



C:Accession: C54471  
R:Garcia, M.L.; Garcia-Calvo, M.; Hidalgo, P.; Lee, A.; MacKinnon, R.  
Biochemistry 33, 6834-6839, 1994  
A:Title: Purification and characterization of three inhibitors of voltage-dependent K<sup>+</sup>  
A:Reference number: A54471; MUID:94263998; PMID:8204618  
A:Accession: C54471  
A:Molecule type: protein  
A:Residues: 1-38 <GAR>  
C:Superfamily: kallotoxin  
C:Keywords: neurotoxin; potassium channel inhibitor; venom  
F:8-28,14-33,18-35/Disulfide bonds: #status predicted

Query Match 12.0%; Score 47; DB 2; Length 38;  
Best Local Similarity 37.8%; Pred. No. 3e+02;  
Matches 14; Conservative 3; Mismatches 16; Indels 4; Gaps 2;

QY 30 PNWVP---KPLCIKICAPGCVCRLGLRNKKVCPVR 63  
| | | | | : | | | | : | | | | : | | | | :  
Db 3 PINVPGTSPQCIRPKCKDAGM-RFGKCMNRKCHCTPK 38

RESULT 6  
B54471  
agitoxin 2 - scorpion (Leiurus quinquestratus)  
N:Alternate names: AGTX-2  
C:Species: Leiurus quinquestratus hebraeus  
C>Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 23-May-1997  
C:Accession: B54471  
R:Garcia, M.L.; Garcia-Calvo, M.; Hidalgo, P.; Lee, A.; MacKinnon, R.  
Biochemistry 33, 6834-6839, 1994  
A:Title: Purification and characterization of three inhibitors of voltage-dependent K<sup>+</sup>  
A:Reference number: A54471; MUID:94263998; PMID:8204618  
A:Accession: B54471  
A:Molecule type: protein  
A:Residues: 1-38 <GAR>  
C:Superfamily: kallotoxin  
C:Keywords: neurotoxin; potassium channel inhibitor; venom  
F:8-28,14-33,18-35/Disulfide bonds: #status predicted

Query Match 11.9%; Score 46.5; DB 2; Length 38;  
Best Local Similarity 39.3%; Pred. No. 3.4e+02;  
Matches 11; Conservative 3; Mismatches 13; Indels 1; Gaps 1;

QY 36 PLCTIKICAPGCVCRLGLRNKKVCPVR 63  
| | | | | : | | | | : | | | | : | | | | :  
Db 12 PQCIKPCCKDAGM-RFGKCMNRKCHCTPK 38

RESULT 7  
S18174  
metallothionein - common bobwhite (fragment)  
C:Species: Colinus virginianus (common bobwhite)  
C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 20-Aug-1999  
C:Accession: S33379; S18174  
R:Shartzel, K.L.; Kage, K.; Sobieski, R.J.; Andrews, G.K.  
J. Mol. Evol. 36, 255-262, 1993  
A:Title: Evolution of avian metallothionein: DNA sequence analyses of the turkey meta  
A:Reference number: S33378; MUID:93247066; PMID:8483164  
A:Accession: S33379  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-43 <SHA>  
A:Cross-references: EMBL:X62512; NID:g62751; PIDN:CAA44371.1; PID:g62752  
C:Superfamily: metallothionein

Query Match 11.0%; Score 46.5; DB 2; Length 43;  
Best Local Similarity 31.0%; Pred. No. 3.7e+02;  
Matches 13; Conservative 5; Mismatches 17; Indels 7; Gaps 3;

QY 8 GRGKCPSNEIFSRCDGRFCPNVVPKPLCIKICAPGCVCYR 49  
| | | | | : | | | | : | | | | : | | | | :  
Db 4 GSCKKN-----CRCS-CRKSCCCCPAG--CNNCVKGCVCK 38

---

QY 8 GRGKCPSNEIFSRCDGRFCPNVVPKPLCIKICAPGCVCYR 49  
| | | | | : | | | | : | | | | : | | | | :  
Db 4 GSCKKN-----CRCS-CRKSCCCCPAG--CNNCAKGCVCCK 38

RESULT 3  
A42040  
kallotoxin 1 [validated] - scorpion (Androctonus mauretanicus)  
N:Alternate names: KTX1  
C:Species: Androctonus mauretanicus  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 15-Sep-2000  
C:Accession: A42040  
R:Crest, M.; Jaquet, G.; Gola, M.; Zerrouk, H.; Benslimane, A.; Rochat, H.; Mansuelle,  
J. Biol. Chem. 267, 1640-1647, 1992  
A:Title: Kallotoxin, a novel peptidyl inhibitor of neuronal BK-type Ca(2+)-activated K<sup>+</sup>  
A:Reference number: A42040; MUID:92112881; PMID:1730708  
A:Accession: A42040  
A:Molecule type: protein  
A:Residues: 1-37 <CRE>  
A>Note: sequence extracted from NCBI backbone (NCBIP:76008); conformation and disulfide  
R:Fernandez, I.; Roml, R.; Szendefi, S.; Martin-eauclair, M.F.; Rochat, H.; Van Rietsch  
submitted to the Brookhaven Protein Data Bank, June 1994  
A:Reference number: A67062; PDB:1KTX  
A:Contents: annotation; conformation by (1)H-NMR, residues 1-37  
C:Comment: This toxin inhibits neuronal BK-type calcium-activated potassium channels.  
C:Superfamily: kallotoxin  
C:Keywords: neurotoxin; potassium channel inhibitor; venom  
F:8-28,14-33,18-35/Disulfide bonds: #status experimental

Query Match 12.3%; Score 48; DB 1; Length 37;  
Best Local Similarity 29.8%; Pred. No. 2.4e+02;  
Matches 14; Conservative 4; Mismatches 17; Indels 12; Gaps 2;

QY 16 EIFSRCDGRFCPNVVPKPLCIKICAPGCVCRLGLRNKKVCPV 62  
| | | | | : | | | | : | | | | : | | | | :  
Db 3 EINVKSG-----SQCLKPCCKDAGM-RFGKCMNRKCHCTP 37

RESULT 4  
A47753  
beta-defensin-10 - bovine  
N:Alternate names: peptide BNBD-10  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 24-Feb-1994 #sequence\_revision 22-Apr-1995 #text\_change 25-Oct-1996  
C:Accession: A47753  
R:Seastedt, M.E.; Tang, Y.Q.; Morris, W.L.; McGuire, P.A.; Novotny, M.J.; Smith, W.; Hens  
J. Biol. Chem. 268, 6641-6648, 1993  
A:Title: Purification, primary structures, and antibacterial activities of beta-defensin  
A:Reference number: A45495; MUID:93203264; PMID:8454635  
A:Accession: A47753  
A:Molecule type: protein  
A:Residues: 1-40 <SE>  
A>Note: sequence modified after extraction from NCBI backbone  
C:Keywords: antibacterial; disulfide bond; pyroglutamic acid  
F:1-40/Product: beta-defensin-10 status experimental <MAL>  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:9-38,16-31,21-39/Disulfide bonds: #status predicted

Query Match 12.3%; Score 48; DB 2; Length 40;  
Best Local Similarity 29.4%; Pred. No. 2.5e+02;  
Matches 10; Conservative 7; Mismatches 13; Indels 1;

QY 7 GGRGKCPSNEIFSRCDGRFCPNVVPKPLCIK 40  
| | | | | : | | | | : | | | | : | | | | :  
Db 11 GNRGIC----LLNRCPGRMRQIGTCIAPRVAKCCR 40

RESULT 5  
C54471  
agitoxin 3 - scorpion (Leiurus quinquestratus)  
N:Alternate names: AGTX-3  
C:Species: Leiurus quinquestratus hebraeus  
C>Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 23-May-1997

## RESULT 8

S43283  
gallinacin - chicken (fragment)  
C.Species: Gallus gallus (chicken)  
C.Date: 20-Oct-1994 #sequence\_revision 10-Nov-1995 #text\_change 07-May-1999  
C.Accession: S43283  
R.Harwig, S.S.L.; Swiderski, K.M.; Kokryakov, V.N.; Tan, L.; Lee, T.D.; Panyutich, E.A.;  
FEBS Lett. 342, 281-285, 1994  
A.Title: Gallinacins: cysteine-rich antimicrobial peptides of chicken leukocytes.  
A.Reference number: S43282; MUID: 94200386; PMID: 8150085  
A.Accession: S43283  
A.Status: preliminary  
A.Molecule type: protein  
A.Residues: 1-39 <VAR>



A:Reference number: A32038; MUID:89123282; PMID:2914898

A:Accession: B32038  
A:Molecule type: protein  
A:Residues: 1-37 <SKI>  
C:Comment: mu-agatoxins cause paralysis in insects by activating neuronal sodium channel  
C:Superfamily: curatotoxin  
C:Keywords: amidated carboxyl end; myotoxin; venom  
F:2-17,9-23,18-33,25-31/Disulfide bonds: #status predicted  
F:37/Modified site: amidated carboxyl end (Ser) #status predicted

Query Match 10.2%; Score 40; DB 1; Length 37;

Best Local Similarity 28.1%; Pred. No. 1.4e+03;

Matches 9; Conservative 3; Mismatches 2; Indels 18; Gaps 2;

Qy 21 CDG---RCQRFCPNVVKPLCIKICAPGVCVR 49

Db 18 CDGLYCSRSY-----PGCMCR 34

# RESULT 18

NTSRPM

C:Species: Androctonus mauretanicus

C:Date: 15-Oct-1982 #sequence\_revision 15-Oct-1982 #text\_change 23-Aug-1996

C:Accession: A01758

ToxiCon 23, 113-125, 1985

A:Title: Characterization of ten proteins from the venom of the Moroccan scorpion Androctonus

A:Reference number: A94318; MUID:85193276; PMID:3992595

A:Accession: A01758

A:Molecule type: protein

A:Residues: 1-35 <ROS>

C:Superfamily: scorpion neurotoxin

C:Keywords: neurotoxin; venom

F:1-18,4-25,15-30,19-32/Disulfide bonds: #status predicted

Query Match 10.1%; Score 39.5; DB 1; Length 35;

Best Local Similarity 31.4%; Pred. No. 1.5e+03;

Matches 11; Conservative 2; Mismatches 3; Indels 19; Gaps 1;

Qy 7 GGRGKCPNIFSRCDGRCQFCPNVVKPLCIKI 41

Db 20 GGRGKC-----VGPOCLCNRI 35

# RESULT 19

S39034

C:Species: Vitis sp. (grape)

C:Date: 25-Dec-1994 #sequence\_revision 19-Apr-1996 #text\_change 17-Mar-1999

C:Accession: S39034

R:Coutos-Thievenot, P.; Jouenne, T.; Maes, O.; Guerbette, F.; Grosbois, M.; le Caer, J.P.

A:Title: Four 9-kDa proteins excreted by somatic embryos of grapevine are isoforms of 11

A:Reference number: S39034; MUID:94039144; PMID:8223644

A:Accession: S39034

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-38 <PGR>

C:Superfamily: phospholipid transfer protein

Query Match 10.1%; Score 39.5; DB 2; Length 38;

Best Local Similarity 55.6%; Pred. No. 1.6e+03;

Matches 10; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

Qy 50 LGYLRNKKVKVPRSKCG 67

Db 15 LGYLRNKKVAVPPGSSCG 31

# RESULT 20

A34471

agitoxin 1 - scorpion (Leiurus quinquestriatus)

N:Alternate names: ACTX-1

C:Species: Leiurus quinquestriatus hebraeus

C:Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 23-May-1997

C:Accession: A54471

R:Garcia, M.L.; Garcia-Calvo, M.; Hidalgo, P.; Lee, A.; MacKinnon, R.

Biochemistry 33, 6834-6839, 1994

A:Title: Purification and characterization of three inhibitors of voltage-dependent K

A:Reference number: A54471; MUID:94263998; PMID:8204618

A:Accession: A54471

A:Molecule type: protein

A:Residues: 1-38 <GAR>

C:Superfamily: kalitoxin

C:Keywords: neurotoxin; potassium channel inhibitor; venom

F:8-28,14-33,18-35/Disulfide bonds: #status predicted

Query Match 10.0%; Score 39; DB 2; Length 38;

Best Local Similarity 27.3%; Pred. No. 1.8e+03;

Matches 12; Conservative 4; Mismatches 16; Indels 12; Gaps 2;

Qy 20 RCDGRCQFCPNVVKPLCIKICAPGVCVRLGYLRNKKVKVPR 63

Db 7 KCTG-----SPQCLPKCKDAGM-RFGKINGKCHCTPK 38

# RESULT 21

S43282

C:Species: Gallus gallus (chicken)

C:Date: 20-Oct-1994 #sequence\_revision 10-Nov-1995 #text\_change 07-May-1999

C:Accession: S43282

R:Harwig, S.S.L.; Swiderek, K.M.; Kokryakov, V.N.; Tan, L.; Lee, T.D.; Panyutich, E.A.

FEBS Lett. 342, 281-285, 1994

A:Title: Gallinacins: cysteine-rich antimicrobial peptides of chicken leukocytes.

A:Reference number: S43282; MUID:94200386; PMID:8150085

A:Accession: S43282

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-39 <HAR>

Query Match 10.0%; Score 39; DB 2; Length 39;

Best Local Similarity 36.7%; Pred. No. 1.9e+03;

Matches 11; Conservative 2; Mismatches 7; Indels 10; Gaps 2;

Qy 11 KCPNIFSRCDGRCQFCPNVVKPLCIK 40

Db 17 KCPYLTLS---GKCSRF-----HLCK 36

# RESULT 22

B34923

C:Species: Agelenopsis aperta

C:Date: 27-Jul-1990 #sequence\_revision 27-Jul-1990 #text\_change 30-Sep-1993

C:Accession: B34923

R:Adams, M.E.; Bindokas, V.P.; Hasegawa, L.; Venema, V.J.

J. Biol. Chem. 265, 861-867, 1990

A:Title: omega-Agatoxins: novel calcium channel antagonists of two subtypes from funn

A:Reference number: A34923; MUID:90110147; PMID:2295621

A:Accession: B34923

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-36 <ADA>

Query Match 9.8%; Score 38.5; DB 2; Length 36;

Best Local Similarity 34.5%; Pred. No. 2e+03;

Matches 10; Conservative 4; Mismatches 10; Indels 5; Gaps 2;

Qy 5 GLGGRGKCPNIFSRCDG---RCQRFCP 30

Db 3 GLPEGAECDDGNECNCKAGAWIKCR--CP 29

# RESULT 23

A44336

neurotoxin Tx3-1 - spider (Phoneutria nigriventer)

C:Species: Phoneutria nigriventer

C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Aug-1998

C:Accession: A44336

R:Cordellero M do, N.; de Figueiredo, S.G.; Valentim A do, C.; Diniz, C.R.; von Eickstedt,

Toxicon 31, 35-42, 1993

A:Title: Purification and amino acid sequences of six Tx3 type neurotoxins from the venom

A:Reference number: A44336; MUID:93190315; PMID:8446961

A:Accession: A44336

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-40 &lt;COR&gt;

A:Experimental source: venom

A:Note: sequence extracted from NCBI backbone (NCBIP:126772)

C:Superfamily: curatatoxin

C:Keywords: neurotoxin

Query Match 9.8%; Score 38.5; DB 2; Length 40;

Best Local Similarity 24.2%; Pred. No. 2.1e+03;

Matches 8; Conservative 5; Mismatches 17; Indels 3; Gaps 1;

QY 17 IFSRCGRCQRCPCPNVFKPLCIKICAPGCVCVR 49

:: || : | : | : | : | : | : |

DB 6 VYERGKGYKRCCEE---RPCKNCIVMDNCTCK 35

RESULT 24

MARSW

myotoxin a 6 - prairie rattlesnake

C:Species: Crotalus viridis viridis (prairie rattlesnake)

C:Date: 30-Apr-1979 #sequence\_revision 30-Apr-1979 #text\_change 07-May-1999

C:Accession: A01736; S27115; A39560; JC5817

R:FOX, J.W.; Elzinga, M.; Tu, A.T.

Biochemistry 18, 678-684, 1979

A:Title: Amino acid sequence and disulfide bond assignment of myotoxin a isolated from t

A:Reference number: A01736; MUID:79124714; PMID:570412

A:Accession: A01736

A:Molecule type: protein

A:Residues: 1-42 &lt;FOX&gt;

R:Baker, B.; Utaisincharoen, P.; Tu, A.T.

Arch. Biochem. Biophys. 298, 325-331, 1992

A:Title: Structure-function relationship of myotoxin a using peptide fragments.

A:Reference number: S27115; MUID:93037463; PMID:1329655

A:Accession: S27115

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-42 &lt;BAK&gt;

R:Aird, S.D.; Kruggel, W.G.; Kaiser, I.I.

Toxicon 29, 265-268, 1991

A:Title: Multiple myotoxin sequences from the venom of a single prairie rattlesnake (Cro

A:Reference number: A39560; MUID:91263105; PMID:2048143

A:Accession: A39560

A:Molecule type: protein

A:Residues: 1-42 &lt;AIR&gt;

R:Nedelkov, D.; O'Keefe, M.P.; Chapman, T.L.; Bieber, A.L.

Biochem. Biophys. Res. Commun. 241, 525-529, 1997

A:Title: The role of Pro20 in the isomerization of myotoxin a from Crotalus viridis viri

A:Reference number: JC5817; MUID:98086384; PMID:9425304

A:Accession: JC5817

A:Molecule type: protein

A:Residues: 1-42 &lt;NED&gt;

C:Comment: This protein inhibits Ca2+ uptake into the sarcoplasmic reticulum by binding

C:Superfamily: crotamine

C:Keywords: myotoxin

F; 4-36, 11-30, 18-37/Disulfide bonds: #status experimental

Query Match

9.7%; Score 38; DB 1; Length 42;

Best Local Similarity 36.8%; Pred. No. 2.5e+03;

Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 46 CVCRGLGYLRNKKKVCVPRS 64

: | : | : | : | : | : |

Db 4 CHKKGGHCFPKKICIPPS 22

RESULT 25

C41711

defensin C - beetle (Zophobas atratus)

C:Species: Zophobas atratus

C:Date: 24-Jul-1992 #sequence\_revision 24-Jul-1992 #text\_change 03-May-1996

C:Accession: C41711

R:Bulet, P.; Cociancich, S.; Dimarcq, J.L.; Lambert, J.; Reichhart, J.M.; Hoffmann, D

J. Biol. Chem. 266, 24520-24525, 1991

A:Title: Insect immunity. Isolation from a coleopteran insect of a novel inducible an

A:Reference number: A41711; MUID:92105112; PMID:1761552

A:Accession: C41711

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-43 &lt;BUL&gt;

C:Superfamily: sapecin precursor

Query Match 9.7%; Score 38; DB 2; Length 43;

Best Local Similarity 41.7%; Pred. No. 2.5e+03;

Matches 10; Conservative 1; Mismatches 11; Indels 2; Gaps 1;

QY 42 CAPCCVC--RLGYLRNKKKVCVPR 63

| : | : | : | : | : |

DB 20 CGAHCIALGRTGGYCNKSKVCVCR 43

Search completed: February 27, 2003, 09:04:45

Job time : 17 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 27, 2003, 09:03:04 ; Search time 13 seconds  
(without alignments)  
213.763 Million cell updates/sec

Title: US-09-506-978-1  
Perfect score: 391  
Sequence: 1 GGFGLGGRKCPNSIFSR.....CRLGLRNKKKVCVPSKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 1266

Minimum DB seq length: 32  
Maximum DB seq length: 45

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50.5	12.9	43	1 MTA_COLVI	P27086 colinus vir
2	50	12.8	38	1 SKL1 ANDMA	P24662 androctonus
3	48	12.3	40	1 BD01_BOVIN	P46168 bos taurus
4	47	12.0	38	1 SCA3_LEIQH	P46112 leiurus qui
5	46.5	11.9	38	1 SCA2_LEIQH	P27087 colinus vir
6	46.5	11.9	43	1 MTB_COLVI	P56686 androctonus
7	46	11.8	37	1 DEF4_ANDAU	P80389 gallus gall
8	45	11.5	39	1 AMPL_CHICK	P81502 aedes aegypt
9	45	11.5	40	1 DEF8_AEDAE	Q01157 lycopersico
10	45	11.5	43	1 GRW1_LICES	P41965 leiurus qui
11	44	11.3	38	1 DEF4_LEIQH	P82226 hadronyche
12	42.5	10.9	36	1 TXJB_HADVE	P07853 cucurbita m
13	42	10.7	32	1 ITR4_CUCMA	P82409 momordica c
14	42	10.7	34	1 ITR2_MOMCO	P11956 drosophila
15	41.5	10.6	43	1 MT2_DROME	P82408 momordica c
16	40.5	10.4	34	1 ITR1_MOMCO	P80571 mytilus gal
17	40.5	10.4	38	1 DEF1_MITGA	P11058 agelenopsis
18	40	10.2	37	1 TXM2_AGEAP	P01498 androctonus
19	39.5	10.1	35	1 SCXP_ANDMA	P82227 hadronyche
20	39.5	10.1	36	1 TXJA_HADVE	P80275 vitis sp. (
21	39.5	10.1	38	1 NLT1_VITXS	P33556 vitis sp. (
22	39.5	10.1	38	1 NLT2_VITXS	P46110 leiurus qui
23	39	10.0	38	1 SCAL_LEIQH	P46157 gallus gall
24	39	10.0	39	1 GLI2_CHICK	P58606 peirates tu
25	38.5	9.8	34	1 PTU1_PEITU	P29187 stoichactis
26	38.5	9.8	35	1 TXKS_STOHE	P15970 agelenopsis
27	38.5	9.8	36	1 TX1B_AGEAP	P82228 hadronyche
28	38.5	9.8	37	1 TXJC_HADVE	P01476 crotalus vi
29	38	9.7	42	1 MYX1_CROVV	P80359 hevea bras
30	38	9.7	45	1 HEVP_HEVBR	P82852 hadronyche
31	38	9.7	45	1 TXO2_HADVE	P55928 pandinus im
32	37.5	9.6	35	1 SCXB_PANIM	Q9x7a2 mycobacteri
33	37.5	9.6	37	1 RL36_MYCLE	

34 37 9.5 32 1 ITR2\_CUCSA  
35 37 9.5 32 1 ITR3\_CUCPE  
36 37 9.5 40 1 BD02\_BOVIN  
37 37 9.5 42 1 CXS3\_CONGE  
38 37 9.5 43 1 DEFA\_ZOPAT  
39 36.5 9.3 37 1 RL36\_BACSU  
40 36.5 9.3 37 1 RL36\_MYCTU  
41 36.5 9.3 38 1 SCX8\_LEIQH  
42 36.5 9.3 40 1 MT1\_DROME  
43 36.5 9.3 40 1 MT1\_DROSI  
44 36 9.2 33 1 Y656\_TREPA  
45 36 9.2 37 1 TXM4\_AGEAP  
46 36 9.2 37 1 TXM5\_AGEAP  
47 36 9.2 38 1 BD01\_BOVIN  
48 36 9.2 40 1 BD07\_BOVIN  
49 36 9.2 42 1 MYXC\_CRODU  
50 35.5 9.1 38 1 DEFI\_AESCY

## ALIGNMENTS

RESULT 1  
MTA\_COLVI STANDARD; PRT; 43 AA.  
AC P27086;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Metallothionein A (MTA) (Fragment).  
OS Colinus virginianus (Bobwhite quail) (Common bobwhite).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
OC Odontophoridae; Colinus.  
OX NCBI\_Taxid=9014;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Liver;  
RC MEDLINE=93247066; PubMed=8483164;  
RA Shartzer K.L., Kage K., Sobieski R.J., Andrews G.K.;  
RT "Evolution of avian metallothionein: DNA sequence analyses of the  
turkey metallothionein gene and metallothionein cDNAs from pheasant  
J. Mol. Evol. 36:255-262(1993).  
CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE  
RESIDUES THAT BIND VARIOUS HEAVY METALS.  
CC -!- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:  
FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA  
DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11  
CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE  
BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.  
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X62511; CAA44370.1; -  
DR PIR; S18173; S18173.  
DR PIR; S33378; S33378.  
DR PIR; S33380; S33380.  
DR PIR; S33382; S33382.  
DR HSP; P28184; IJ19.  
DR InterPro; IPR003019; Metallthion.  
DR InterPro; IPR000006; Metallthion\_vert.  
DR Pfam; PF00131; metalthio; 1.  
DR PROSITE; PS00203; METALLOTHIONEIN\_VRT; PARTIAL.  
KW Metal-binding; Metal-thiolate cluster.  
FT NON\_TER 1

1:



```

RESULT 4
SCA2_LEIQH
ID SCA2_LEIQH STANDARD; PRT; 38 AA.
AC P46112;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Agitoxin 3 (AGTX-3).
OS Leiurus quinquestriatus hebraeus (Scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Leiurus.
OX NCBI_TaxID=6884;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=94263998; PubMed=8204618;
RA Garcia M.L., Garcia-Calvo M., Hidalgo P., Lee A., Mackinnon R.;
RT "Purification and characterization of three inhibitors of voltage-
RT dependent K+ channels from Leiurus quinquestriatus var. hebraeus
RT venom.";
RL Biochemistry 33:6834-6839(1994).
CC -!- FUNCTION: POTENT INHIBITOR OF SHAKER POTASSIUM CHANNELS AS WELL AS
CC THE MAMMALIAN HOMOLOGS OF SHAKER.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SHORT SCORPION TOXIN FAMILY. POTASSIUM
CC CHANNEL INHIBITORS SUBFAMILY.
DR HSSP; P46111; IAGT.
DR InterPro: IPR001947; Scorpion_toxins.
DR Pfam; PF00451; toxin_2; 1.
DR ProDom; PD003586; Scorpion_toxins; 1.
DR PROSITE; PS01138; SCORP_SHORT_TOXIN; 1.
KW Neurotoxin; Potassium channel inhibitor.
FT DISULFID 8 28 BY SIMILARITY.
FT DISULFID 14 33 BY SIMILARITY.
FT DISULFID 18 35 BY SIMILARITY.
FT SITE 26 33 INTERACTION WITH CA(2+)-ACTIVATED K(+)
FT SITE CHANNELS (POTENTIAL).
SQ SEQUENCE 38 AA; 4107 MW; E5945513F87B51CE CRC64;

Query Match 12.0%; Score 47; DB 1; Length 38;
Best Local Similarity 37.8%; Pred. No. 68;
Matches 14; Conservative 3; Mismatches 16; Indels 4; Gaps 2;

Qy 30 PNVPVPG---PLCICKICAPGCVRLGNKVKVCP 63
Db 3 PINVPTGSPQCKPKCKDAGM-RFGKCMNRKCHCTPK 38

RESULT 5
SCA2_LEIQH
ID SCA2_LEIQH STANDARD; PRT; 38 AA.
AC P46111;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Agitoxin 2 (AGTX-2).
OS Leiurus quinquestriatus hebraeus (Scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Leiurus.
OX NCBI_TaxID=6884;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=94263998; PubMed=8204618;
RA Garcia M.L., Garcia-Calvo M., Hidalgo P., Lee A., Mackinnon R.;
RT "Purification and characterization of three inhibitors of voltage-
RT dependent K+ channels from Leiurus quinquestriatus var. hebraeus
RT venom.";
RL Biochemistry 33:6834-6839(1994).
RN [2]
RP STRUCTURE BY NMR.

```

```

RX MEDLINE=96060078; PubMed=8520473;
RA Krezel A.M., Kasibhatla C., Hidalgo P., Mackinnon R., Wagner G.;
RT "Solution structure of the potassium channel inhibitor agitoxin 2:
RT caliper for probing channel geometry.";
RL Protein Sci. 4:1478-1489(1995).
CC -!- FUNCTION: POTENT INHIBITOR OF SHAKER POTASSIUM CHANNELS AS WELL AS
CC THE MAMMALIAN HOMOLOGS OF SHAKER.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SHORT SCORPION TOXIN FAMILY. POTASSIUM
CC CHANNEL INHIBITORS SUBFAMILY.
DR PDB; IAGT; 10-JUL-95.
DR InterPro: IPR001947; Scorpion_toxins.
DR Pfam; PF00451; toxin_2; 1.
DR ProDom; PD003586; Scorpion_toxins; 1.
DR PROSITE; PS01138; SCORP_SHORT_TOXIN; 1.
KW Neurotoxin; Potassium channel inhibitor; 3D-structure.
FT DISULFID 8 28 BY SIMILARITY.
FT DISULFID 14 33 BY SIMILARITY.
FT DISULFID 18 35 BY SIMILARITY.
FT SITE 26 33 INTERACTION WITH CA(2+)-ACTIVATED K(+)
FT SITE CHANNELS (POTENTIAL).
SQ SEQUENCE 38 AA; 4097 MW; C8BB8513F87B51CD CRC64;

Query Match 11.9%; Score 46.5; DB 1; Length 38;
Best Local Similarity 39.3%; Pred. No. 77;
Matches 11; Conservative 3; Mismatches 13; Indels 1; Gaps 1;

Qy 36 PLCICKICAPGCVRLGNKVKVCP 63
Db 12 PQCKPKCKDAGM-RFGKCMNRKCHCTPK 38

RESULT 6
MTB_COLVI
ID MTB_COLVI STANDARD; PRT; 43 AA.
AC P27087;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metallothionein B (MTB) (Fragment).
OS Colinus virginianus (Bobwhite quail) (Common bobwhite).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Odontophorinae; Colinus.
OX NCBI_TaxID=9014;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93247066; PubMed=8483164;
RA Shartzer K.L., Kage K., Sobieski R.J., Andrews G.K.;
RT "Evolution of avian metallothionein: DNA sequence analyses of the
RT turkey metallothionein gene and metallothionein cDNAs from pheasant
RT and quail.";
RL J. Mol. Evol. 36:255-262(1993).
CC -!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE
CC RESIDUES THAT BIND VARIOUS HEAVY METALS.
CC -!- DOMAIN: CLASS I METALLOTHIONEINS CONTAIN 2 METAL-BINDING DOMAINS:
CC FOUR DIVALENT IONS ARE CHELATED WITHIN CLUSTER A OF THE ALPHA
CC DOMAIN AND ARE COORDINATED VIA CYSTEINYL THIOLATE BRIDGES TO 11
CC CYSTEINE LIGANDS. CLUSTER B, THE CORRESPONDING REGION WITHIN THE
CC BETA DOMAIN, CAN LIGATE THREE DIVALENT IONS TO 9 CYSTEINES.
CC -!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X62512; CAA44371.1; -
CC PIR; S18174; S18174.

```

Thu Feb 27 09:12:57 2003

us-09-506-978-1.slim.rsp

PIR: S33379; S33379.  
HSP: P04355; 1MRT.  
InterPro: IPR003019; Metallthion..  
InterPro: IPR000006; Metallthion.vert.  
Pfam: PF00131; metalthio; 1.  
PROSITE: PS00203; METALLOTHIONEIN\_VRT; PARTIAL.  
KW Metal-binding; Metal-thiolate cluster.  
FT NON\_TER 1 1  
FT DOMAIN <1 16 BETA.  
FT METAL 17 >43 ALPHA.  
FT METAL 2 2 CLUSTER B.  
FT METAL 6 6 CLUSTER B.  
FT METAL 8 8 CLUSTER B.  
FT METAL 11 11 CLUSTER B.  
FT METAL 13 13 CLUSTER B.  
FT METAL 16 16 CLUSTER B.  
FT METAL 20 20 CLUSTER A.  
FT METAL 21 21 CLUSTER A.  
FT METAL 23 23 CLUSTER A.  
FT METAL 24 24 CLUSTER A.  
FT METAL 28 28 CLUSTER A.  
FT METAL 31 31 CLUSTER A.  
FT METAL 35 35 CLUSTER A.  
FT METAL 37 37 CLUSTER A.  
FT NON\_TER 43 43  
SQ SEQUENCE 43 AA: 4429 MW; 1612EB40EE6EB875 CRC64;  
Query Match 11.9%; Score 46.5; DB 1; Length 43;  
Best Local Similarity 31.0%; Pred. No. 84;  
Matches 13; Conservative 5; Mismatches 17; Indels 7; Gaps 3;  
QY 8 GRGKCPNFEIFSRCDGRCPNVPKPLCIKICAPCCVCR 49  
DB 4 GSKCKN-----CRKRS-CRKSCCCCPAG--CNCVKGCVCCK 38  
RESULT 7  
DEF4\_ANDAU STANDARD; PRT; 37 AA.  
ID DEF4\_ANDAU  
AC P56686; P81618;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE 4 kDa defensin.  
OS Androctonus australis hector (Sahara scorpion).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
OC Butyridae; Butyridae; Androctonus.  
OX NCBI\_TaxID=70175;  
RN [1]  
RP SEQUENCE, AND CHARACTERIZATION.  
RC TISSUE=Hemolymph;  
RX MEDLINE=97094646; PubMed=8939880;  
RA Ehret-Sabatier L., Loew D., Goyffon M., Fehlbaum P., Hoffmann J.A.,  
RA van Dorsselaer A., Bulet P.;  
RT "Characterization of novel cysteine-rich antimicrobial peptides from  
scorpion blood";  
RL J. Biol. Chem. 271:29537-29544 (1996).  
CC -1- FUNCTION: ACTIVE AGAINST GRAM-POSITIVE BACTERIA.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- MASS SPECTROMETRY: MW=4206.8; METHOD=Electrospray.  
CC -1- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.  
DR HSP; P10891; IICA.  
DR InterPro: IPR001542; Defensin\_anpod.  
DR Pfam: PF01097; Arthro\_defensin; 1.  
DR PROSITE: PS00425; ARTHROPOD\_DEFENSINS; 1.  
KW Antibiotic.  
FT DISULFID 4 25 BY SIMILARITY.  
FT DISULFID 11 33 BY SIMILARITY.  
FT DISULFID 15 35 BY SIMILARITY.  
SQ SEQUENCE 37 AA: 4212 MW; AB1363ECE3FB84C1 CRC64;  
Query Match 11.8%; Score 46; DB 1; Length 37;  
Best Local Similarity 27.1%; Pred. No. 85;

Matches 13; Conservative 5; Mismatches 16; Indels 14; Gaps 3;  
QY 2 GFGLGGRGKCPNFEIFSRCDGRCPNVPK-PLCIKICAPGCVC 48  
DB 1 GFG-----CPFNQ-----GACHRHCRSIRRRGGYAGLKFQKQTC 35  
RESULT 8  
AMPI\_CHICK STANDARD; PRT; 39 AA.  
ID AMPI\_CHICK  
AC P80389;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Antimicrobial peptide, CHPI (Chicken heterophil peptide 1).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95053386; PubMed=7964174;  
RA Evans E.W., Beach G.G., Wunderlich J., Harmon B.G.;  
RT "Isolation of antimicrobial peptides from avian heterophils";  
RL J. Leukoc. Biol. 56:661-665 (1994).  
CC -1- FUNCTION: BACTERICIDAL ACTIVITY; INHIBITS STAPHYLOCOCCUS AUREUS  
CC AND ESCHERICHIA COLI.  
KW Antibiotic.  
FT DISULFID 6 28 BY SIMILARITY.  
FT DISULFID 13 34 BY SIMILARITY.  
FT DISULFID 18 35 BY SIMILARITY.  
SQ SEQUENCE 39 AA; 4480 MW; 5DDF1051693D254E CRC64;  
Query Match 11.5%; Score 45; DB 1; Length 39;  
Best Local Similarity 40.0%; Pred. No. 11e+02;  
Matches 12; Conservative 2; Mismatches 6; Indels 10; Gaps 2;

QY 11 KCPSNEIFSRCDGRCPNVPKPLCIK 40  
DB 17 KCPSLTLLIS---GKCSRFV-----LCKK 36  
RESULT 9  
DEFB\_AEDAE STANDARD; PRT; 40 AA.  
ID DEFB\_AEDAE  
AC P81602;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Defensin B.  
OS Aedes aegypti (Yellowfever mosquito).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;  
OC Culicoidae; Aedes.  
OX NCBI\_TaxID=7159;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=Liverpool;  
RX MEDLINE=95360030; PubMed=7633471;  
RA Lowenberger C., Bulet P., Charlet M., Hetru C., Hodgeman B.,  
RA Christensen B.M., Hoffmann J.A.;  
RT "Insect immunity: Isolation of three novel inducible antibacterial  
defensins from the vector mosquito, Aedes aegypti";  
RL Insect Biochem. Mol. Biol. 25:867-873 (1995).  
CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE MOSTLY AGAINST GRAM-POSITIVE  
CC BACTERIA.  
CC -1- INDUCTION: By bacterial infection.  
CC -1- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.  
DR HSP; P10891; IICA.  
DR InterPro: IPR001542; Defensin\_anpod.  
DR InterPro: IPR003614; Knot1.  
DR Pfam: PF01097; Arthro\_defensin; 1.

DD DEF4\_LEIQH STANDARD; PRT; 38 AA.

D DEF4\_LEIQH STANDARD; PRT; 38 AA.

FT	10	22	BY SIMILARITY.
FT	13	14	BY SIMILARITY.

FT	DISULFID	10	22	BY SIMILARITY.
FT	DISULFID	12	14	BY SIMILARITY.

SQ SEQUENCE 36 AA: 3651 MW; D23A442560B89997 CRC64;  
Query Match 10.9%; Score 42.5; DB 1; Length 36;  
Best Local Similarity 26.7%; Pred. No. 1.9e+02;  
Matches 12; Conservative 3; Mismatches 13; Indels 17; Gaps 2;

QY 17 IFSRCDGRCRQFNVVVKPLKICAPGVCVCR-----LGYLRN 55  
DB 2 ICTGADRPCAACCP-----CCGCTSCQGPENGVSYCRN 35

RESULT 13  
ITR4\_CUCMA  
ID ITR4\_CUCMA STANDARD; PRT; 32 AA.  
AC P07853;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 01-MAY-1992 (Rel. 22, Last annotation update)  
DE Trypsin inhibitors IV and III (CMTI-IV and CMTI-III).  
OS Cucurbita maxima (Pumpkin) (Winter squash).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.  
OX NCBI\_TaxID=3661;  
RN [1]  
RP SEQUENCE (CMT-IV).  
RC TISSUE=Seed;  
RX MEDLINE=85149300; PubMed=3977882;  
RA Wiczeorek M., Otlewski J., Cook J., Parks K., Leluk J.,  
RA Wilmowska-Pelc A., Polanowski A., Wilusz T., Laskowski M. Jr.;  
RT "The squash family of serine proteinase inhibitors. Amino acid  
RT sequences and association equilibrium constants of inhibitors from  
RT squash, summer squash, zucchini, and cucumber seeds.";  
RL Biochem. Biophys. Res. Commun. 126:646-652(1985).  
RN [2]  
RP SEQUENCE (CMT-III).  
RC TISSUE=Seed;  
RX MEDLINE=83184077; PubMed=6840699;  
RA Wilusz T., Wiczeorek M., Polanowski A., Denton A., Cook J.,  
RA Laskowski M. Jr.;  
RT "Amino-acid sequence of two trypsin isoinhibitors, ITD I and ITD III  
RT from squash seeds (Cucurbita maxima).";  
RL Hoppe-Seyler's Z. Physiol. Chem. 364:93-95(1983).  
RN [3]  
RP PRELIMINARY SEQUENCE (CMT-III).  
RC TISSUE=Seed;  
RX MEDLINE=82005824; PubMed=7275008;  
RA Nowak K., Slominska A., Polanowski A., Wiczeorek M., Wilusz T.;  
RT "Trypsin inhibitor III from squash seeds (Cucurbita maxima), its  
RT reactive site and proposed amino acid sequence.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 362:1017-1019(1981).  
RN [4]  
RP STRUCTURE BY NMR OF CMT-III.  
RX MEDLINE=92118901; PubMed=1731946;  
RA Krishnamoorthi R., Gong Y., Lin C.-L.S., Vandervelde D.;  
RT "Two-dimensional NMR studies of squash family inhibitors. Sequence-  
RT specific proton assignments and secondary structure of reactive-site  
RT hydrolyzed Cucurbita maxima trypsin inhibitor III.";  
RL Biochemistry 31:898-904(1992).  
CC -1- SIMILARITY: BELONGS TO THE SQUASH FAMILY OF SERINE PROTEASE  
CC INHIBITORS.  
CC PIR; A01313; TIPU3.  
DR PIR; S07156; S07156.  
DR HSSP; P01074; ICTI.  
DR InterPro; IPR000737; Squash.  
DR Pfam; PF00299; squash; 1.  
DR PRINTS; PR00293; SQUASHINHBTR.  
DR ProDom; PD003401; Squash; 1.  
DR SMART; SM00286; PTI; 1.  
DR PROSITE; PS00286; SQUASH\_INHIBITOR; 1.  
KW Serine protease inhibitor.  
FT DISULFID 8 25  
FT DISULFID 15 27  
FT DISULFID 21 33  
SQ SEQUENCE 34 AA: 3477 MW; 8F7D0B4C048BB93A CRC64;

Query Match 10.7%; Score 42; DB 1; Length 34;  
Best Local Similarity 33.3%; Pred. No. 2.1e+02;  
Matches 7; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 29 CPNVVVKPLKICAPGVCVCR 49  
DB 8 CPKILKCRSDSDCPGACICR 28

RESULT 15  
MT2\_DROME  
ID MT2\_DROME STANDARD; PRT; 43 AA.

FT ACT\_SITE .8 9 REACTIVE BOND.  
FT DISULFID 6 23 BY SIMILARITY.  
FT DISULFID 13 25 BY SIMILARITY.  
FT DISULFID 19 31 BY SIMILARITY.  
SQ SEQUENCE 32 AA: 3669 MW; 0F591120B0137512 CRC64;  
Query Match 10.7%; Score 42; DB 1; Length 32;  
Best Local Similarity 34.8%; Pred. No. 2e+02;  
Matches 8; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 26 QRECPNVVVKPLKICAPGVCVCR 48  
DB 3 ERVCPRLMKCKKSDCLAEVCV 25

RESULT 14  
ITR2\_MOMCO  
ID ITR2\_MOMCO STANDARD; PRT; 34 AA.  
AC P82409;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Trypsin inhibitor II (MCOTI-II).  
OS Momordica cochinchinensis (Spiny bitter cucumber).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.  
OX NCBI\_TaxID=3674;  
RN [1]  
RP SEQUENCE, AND MASS SPECTROMETRY.  
RC TISSUE=Seed;  
RX MEDLINE=20263490; PubMed=10801322;  
RA Hernandez J.-F., Gagnon J., Chiche L., Nguyen T.M., Andrieu J.-P.,  
RA Heitz A., Trinh T., Pham T.T.C., Le Nguyen D.;  
RT "Squash trypsin inhibitors from Momordica cochinchinensis exhibit an  
RT atypical macrocyclic structure.";  
RL Biochemistry 39:5722-5730(2000).  
CC -1- FUNCTION: INHIBITS TRYPSIN; PROBABLY PARTICIPATES IN A PLANT  
CC DEFENSE MECHANISM.  
CC -1- PTM: A CYCLIC SUCCINIMIDE MAY BE FORMED BETWEEN ASP-4 AND GLY-5  
CC RESIDUES WHICH CAN BE TRANSFORMED TO A BETA-ASPARTYL BOND. SO  
CC THERE EXISTS THREE ISOFORMS OF MCOTI-II, NORMAL, WITH SUCCINIMIDE  
CC OR WITH A BETA-ASPARTYL BOND.  
CC -1- PTM: THIS IS A CYCLIC PEPTIDE.  
CC -1- MASS SPECTROMETRY: MW=3453; MW ERR=0.2; METHOD=Electrospray.  
CC -1- SIMILARITY: BELONGS TO THE SQUASH FAMILY OF SERINE PROTEASE  
CC INHIBITORS.  
CC HSSP; P01074; IBXJ.  
DR InterPro; IPR000737; Squash.  
DR Pfam; PF00299; squash; 1.  
DR PRINTS; PR00293; SQUASHINHBTR.  
DR ProDom; PD003401; Squash; 1.  
DR SMART; SM00286; PTI; 1.  
DR PROSITE; PS00286; SQUASH\_INHIBITOR; 1.  
KW Serine protease inhibitor.  
FT ACT\_SITE 10 11 REACTIVE BOND.  
FT DISULFID 8 25 BY SIMILARITY.  
FT DISULFID 15 27 BY SIMILARITY.  
FT DISULFID 21 33 BY SIMILARITY.  
SQ SEQUENCE 34 AA: 3477 MW; 8F7D0B4C048BB93A CRC64;



```

KW Serine protease inhibitor.
FT ACT_SITE 10 11 REACTIVE BOND.
FT DISULFID 8 25 BY SIMILARITY.
FT DISULFID 15 27 BY SIMILARITY.
FT DISULFID 21 33 BY SIMILARITY.
SQ SEQUENCE 34 AA; 3505 MW; 8F7D0B4C162C935A CRC64;

Query Match 10.4%; Score 40.5; DB 1; Length 34;
Best Local Similarity 30.0%; Pred. No. 3e+02;
Matches 12; Conservative 2; Mismatches 9; Indels 17; Gaps 3;

QY 10 GKCPNEIFSRCDGRCQFCPNVVPKPLCIKICAPGCVCR 49
DB 6 GVCP--KILQRC--RRSDCPG-----ACICR 28

RESULT 17
DEFL_MYTGA
ID DEFL_MYTGA STANDARD; PRT; 38 AA.
AC P80371;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Defensin MGD-1.
GN FH3.
OS Mytilus galloprovincialis (Mediterranean mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=29158;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=97025339; PubMed=8925841;
RA Hubert F., Noel T., Roch P.;
RT "A member of the arthropod defensin family from edible Mediterranean
mussels (Mytilus galloprovincialis).";
RL Eur. J. Biochem. 240:302-306(1996).
RN [2]
RP ERRATUM.
RA Hubert F., Noel T., Roch P.;
RL Eur. J. Biochem. 240:815-815(1996).
CC -!- SIMILARITY: BELONGS TO THE ARTHROPOD DEFENSIN FAMILY.
DR InterPro: IPR001542; Defensin_anpod.
DR Pfam: PF01097; Arthro_defensin; 1.
DR PROSITE: PS00425; ARTHROPOD_DEFENSINS; 1.
KW Antibiotic.
FT DISULFID 4 25 BY SIMILARITY.
FT DISULFID 10 33 BY SIMILARITY.
FT DISULFID 14 35 BY SIMILARITY.
FT DISULFID 21 38 BY SIMILARITY.
SQ SEQUENCE 38 AA; 4275 MW; 6F1C0CDCF0E69E76 CRC64;

Query Match 10.4%; Score 40.5; DB 1; Length 38;
Best Local Similarity 34.4%; Pred. No. 3e+02;
Matches 11; Conservative 3; Mismatches 7; Indels 11; Gaps 3;

QY 2 GFGLGGRGKCPNS-EIFSR-----DGRQRCFC 29
DB 1 GFG-----CPNNYQCHRHCKSIPIGRGGCYC 25

RESULT 18
TXM2_AGEAP
ID TXM2_AGEAP STANDARD; PRT; 37 AA.
AC P11058;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mu-agatoxin 2.
OS Agelenopsis aperta (Funnel-web spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Agelenidae; Agelenopsis.
OX NCBI_TaxID=6908;

```

```

RN SEQUENCE.
RP MEDLINE=89123282; PubMed=2914898;
RX Skinner W.S., Adams M.E., Quistad G.B., Kataoka H., Cesarin B.J.,
RA Enderlin F.E., Schooley D.A.;
RT "Purification and characterization of two classes of neurotoxins from
the funnel web spider, Agelenopsis aperta.";
RL J. Biol. Chem. 264:2150-2155(1989).
CC -!- FUNCTION: CAUSE IRREVERSIBLE PARALYSIS IN LEPIDOPTEROUS INSECTS
BY MASSIVE TRANSMITTER RELEASE (WHICH IS MEDIATED BY GLUTAMIC ACID
RECEPTORS) FROM PRESYNAPTIC STORES AT NEUROMUSCULAR JUNCTIONS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- SIMILARITY: BELONGS TO THE MU-AGATOXIN FAMILY.
DR PIR: B32038; B32038.
DR HSP; P11060; LEIU.
KW Toxin; Neurotoxin.
FT DISULFID 2 18 BY SIMILARITY.
FT DISULFID 9 23 BY SIMILARITY.
FT DISULFID 17 33 BY SIMILARITY.
FT DISULFID 25 31 BY SIMILARITY.
SQ SEQUENCE 37 AA; 4110 MW; 8900D367B4F096CD CRC64;

Query Match 10.2%; Score 40; DB 1; Length 37;
Best Local Similarity 28.1%; Pred. No. 3.7e+02;
Matches 9; Conservative 3; Mismatches 2; Indels 18; Gaps 2;

QY 21 CDG---RCQFCPNVVPKPLCIKICAPGCVCR 49
DB 18 CDGLYCSCRSY-----PGCMCR 34

RESULT 19
SCXP_ANDMA
ID SCXP_ANDMA STANDARD; PRT; 35 AA.
AC P01498;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neurotoxin P2.
OS Androctonus mauretanicus mauretanicus (Scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Androctonus.
OX NCBI_TaxID=6860;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=85193276; PubMed=3992595;
RA Rosso J.P., Rochat H.;
RT "Characterization of ten proteins from the venom of the Moroccan
scorpion Androctonus mauretanicus mauretanicus, six of which are
toxic to the mouse.";
RL Toxicon 23:113-125(1985).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SHORT SCORPION TOXIN FAMILY.
DR PIR: A01758; NTSRPM.
DR HSP; P15222; 1SIS.
KW Neurotoxin.
FT DISULFID 1 18 BY SIMILARITY.
FT DISULFID 4 25 BY SIMILARITY.
FT DISULFID 15 30 BY SIMILARITY.
FT DISULFID 19 32 BY SIMILARITY.
SQ SEQUENCE 35 AA; 3673 MW; 213B69262289EB5A CRC64;

Query Match 10.1%; Score 39.5; DB 1; Length 35;
Best Local Similarity 31.4%; Pred. No. 4e+02;
Matches 11; Conservative 2; Mismatches 3; Indels 19; Gaps 1;

QY 7 GGRGKCPNSNEIFSRCDGRCQFCPNVVPKPLCIKI 41
DB 20 GGRGK-----VGQCCLNRI 35

```

```

NCBI_L VITISX          STANDARD;          PRT;          38 AA.
ID   NT1_VIT5X          P80275;
AC   P80275;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Nonspecific lipid-transfer protein p1 (LTP p1) (Fragment).
OS   Vitis sp. (Grape).
OC   Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae;
OC   Vitis.
OX   NCBI_TaxID=3604;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=V. vinifera X Berlanchen cv. Rootstock 41B;
RX   MEDLINE=94039144; PubMed=8223644;
RA   Coutos-Thévenot P., Jouenne T., Maes O., Guerbette F., Grosbois M.,
RA   le Caer J.P., Boulay M., Deltoile A., Kader J.-C., Guern J.;
RT   "Four 9-kDa proteins excreted by somatic embryos of grapevine are
RT   isoforms of lipid-transfer proteins.";
RL   Eur. J. Biochem. 217:885-889(1993).
CC   !- FUNCTION: PLANT NONSPECIFIC LIPID-TRANSFER PROTEINS TRANSFER
CC   PHOSPHOLIPIDS AS WELL AS GALACTOLIPIDS ACROSS MEMBRANES. MAY PLAY
CC   A ROLE IN WAX OR CUTIN DEPOSITION IN THE CELL WALLS OF EXPANDING
CC   EPIDERMAL CELLS AND CERTAIN SECRETORY TISSUES.
CC   !- SIMILARITY: BELONGS TO THE PLANT LTP FAMILY.
CC   HSPSP: PL9656; LM2M.
DR   InterPro: IPR000528; plant_LTP.
DR   InterPro: IPR001768; Tryp/amyL_inhbr.
DR   Pfam: PF00234; tryp_alpha_aml; 1.

```

```

RC TISSUE=Venom;
RX MEDLINE=94263998; PubMed=8204518;
RA Garcia M.L., Garcia-Calvo M., Hidalgo P., Lee A., Mackinnon R.;
RT "Purification and characterization of three inhibitors of voltage-
RT dependent K+ channels from Lelurus quinquestriatus var. hebraeus
RT venom.";
RL Biochemistry 33:6834-6839(1994).
CC -!- FUNCTION: POTENT INHIBITOR OF SHAKER POTASSIUM CHANNELS AS WELL AS
CC THE MAMMALIAN HOMOLOGS OF SHAKER.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE SHORT SCORPION TOXIN FAMILY. POTASSIUM
CC CHANNEL INHIBITORS SUBFAMILY.
DR HSP; P46111; IACT.
DR InterPro; IPR001947; Scorpion_toxins.
DR Pfam; PF00451; toxin_2; 1.
DR ProDom; PD003586; Scorpion_toxins; 1.
DR PROSITE; PS01138; SCORP_SHORT_TOXIN; 1.
KW Neurotoxin; Potassium channel inhibitor.
FT DISULFID 8 28 BY SIMILARITY.
FT DISULFID 14 33 BY SIMILARITY.
FT DISULFID 18 35 BY SIMILARITY.
FT SITE 26 33 INTERACTION WITH CA(2+)-ACTIVATED K(+)
FT CHANNELS (POTENTIAL).
FT SEQUENCE 38 AA; 4021 MW; A0951113F87B51CE CRC64;

Query Match 10.0%; Score 39; DB 1; Length 38;
Best Local Similarity 27.3%; Pred. No. 4.8e+02;
Matches 12; Conservative 4; Mismatches 16; Indels 12; Gaps 2;

QY 20 RCDGRCQRCPCPNVWPPLCIKICAPGCVCRGLYLRNKKKVCVPR 63
DB 7 KCTG-----SPOCLPKCKDAGM-RFGKCKNGKCHCTPK 38

RESULT 24
GLL2_CHICK
ID GLL2_CHICK STANDARD; PRT; 39 AA.
AC P46157;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Gallinacin 1 alpha.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RC STRAIN=Cross Broiler-6; TISSUE=Leukocyte;
RX MEDLINE=94200386; PubMed=8150085;
RA Harvig S.S.L., Swiderok K.M., Kokryakov V.N., Tan L., Lee T.D.,
RA Panyutich E.A., Aleshina G.M., Shamova O.V., Lehrer R.I.;
RT "Gallinacins: cysteine-rich antimicrobial peptides of chicken
RT leukocytes.";
RL FEBS Lett. 342:281-285(1994).
CC -!- FUNCTION: HAS BACTERICIDAL ACTIVITY. POTENT AGAINST E. COLI ML-35,
CC L. MONOCYTOGENES EGD AND C. ALBICANS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES.
CC -!- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY.
KW Antibiotic.
FT DISULFID 6 34 BY SIMILARITY.
FT DISULFID 13 28 BY SIMILARITY.
FT DISULFID 18 35 BY SIMILARITY.
FT SEQUENCE 39 AA; 4587 MW; AFAF105C0441F499 CRC64;

Query Match 10.0%; Score 39; DB 1; Length 39;
Best Local Similarity 36.7%; Pred. No. 4.9e+02;
Matches 11; Conservative 2; Mismatches 7; Indels 10; Gaps 2;

QY 11 KCPSNEIFSRCDGRCPCPNVWPPLCIK 40
DB 17 KCPYLTLIS--GKCSR-----HLCKK 36

```

```

RESULT 25
PTUL_PTITU
ID PTUL_PTITU STANDARD; PRT; 34 AA.
AC P58606;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Toxin Ptul.
OS Peirates turpis (Assassin bug).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC Heteroptera; Panheteroptera; Cimicomorpha; Reduviidae; Peiratinae;
OC Peirates.
OX NCBI_TaxID=181095;
RN [1]
RP SEQUENCE, MASS SPECTROMETRY, SYNTHESIS, AND CHARACTERIZATION.
RC TISSUE=Saliva;
RX MEDLINE=21316029; PubMed=11423127;
RA Corzo G., Adachi-Akahane S., Nagao T., Kusui Y., Nakajima T.;
RT "Novel peptides from assassin bugs (Hemiptera: Reduviidae): isolation,
RT chemical and biological characterization.";
RL PEBS Lett. 499:256-261(2001).
RN [2]
RP STRUCTURE BY NMR.
RX MEDLINE=21526409; PubMed=11669615;
RA Bernard C., Corzo G., Mosbah A., Nakajima T., Darbon H.;
RT "Solution structure of Ptul, a toxin from the assassin bug Peirates
RT turpis that blocks the voltage-sensitive calcium channel N-type.";
RL Biochemistry 40:12795-12800(2001).
CC -!- FUNCTION: Binds reversibly and blocks N-type voltage-gated calcium
CC channels.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Produced by the venomous saliva.
CC -!- MASS SPECTROMETRY: MW=3615.1; METHOD=MALDI.
DR PDB; 1I26; 21-NOV-01.
KW Calcium channel inhibitor; Toxin; Neurotoxin; 3D-structure.
FT DISULFID 5 20
FT DISULFID 12 26
FT DISULFID 19 33
FT SEQUENCE 34 AA; 3621 MW; 6858A20A9E6B6FFA CRC64;

Query Match 9.8%; Score 38.5; DB 1; Length 34;
Best Local Similarity 37.9%; Pred. No. 4.9e+02;
Matches 11; Conservative 1; Mismatches 10; Indels 7; Gaps 2;

QY 38 CIKICAPGCVCRGLYLRNKKKVCVPRSKC 66
DB 5 CI---APGAPC----FGTDRPCCNPAWC 26

Search completed: February 27, 2003, 09:05:05
Job time : 15 secs

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 27, 2003, 09:03:04 : Search time 28 Seconds  
(without alignments)  
493.041 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GGFGLGGRKCPSEIFSR.....CRGLGLRNKKVCVPSKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 13368

Minimum DB seq length: 32

Maximum DB seq length: 45

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database :

SPTREMBL\_21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriophage.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query No.	Score	Match	Length	DB	ID	Description
1	62.5	16.0	40	13	Q98TP9		Q98tp9 platichthys
2	62.5	16.0	44	11	Q99KF9		Q99kf9 mus musculus
3	53	13.6	39	13	Q9PVG7		Q9pv7 xiphophorus
4	48.5	12.4	41	4	Q9UDP7		Q9udp7 homo sapien
5	47.5	12.1	37	11	Q9QX87		Q9qx87 rattus norv
6	47	12.0	41	13	Q9PVG6		Q9pv6 xiphophorus
7	45	11.5	36	10	Q02023		Q02023 lycopersico
8	44.5	11.4	37	4	Q9H499		Q9h499 homo sapien
9	44	11.3	40	4	Q8WV73		Q8wv73 homo sapien
10	44	11.3	45	10	Q9LW92		Q9lw92 nicotiana t
11	42.5	10.9	42	5	O18625		O18625 schistosoma
12	42	10.7	39	5	P82380		P82380 stomoxys ca
13	41.5	10.6	43	5	Q9VDN2		Q9vbn2 drosophila
14	41.5	10.6	43	13	Q8UUG4		Q8uug4 gallus gall
15	41.5	10.6	45	10	O81529		O81529 mesembryant
16	40	10.2	33	5	Q17071		Q17071 antheraea p

17	40	10.2	37	5	P83259	P83259 paracoelote
18	40	10.2	45	5	Q9VM82	Q9vm82 drosophila
19	39	10.0	35	10	Q94214	Q94214 oryza sativ
20	39	10.0	42	16	Q8XTX4	Q8xtx4 raistonia s
21	38	9.7	38	10	Q8W2G7	Q8w2g7 oryza sativ
22	37.5	9.6	42	8	Q9MRL7	Q9mrl7 daucus caro
23	37	9.5	33	4	Q9UD12	Q9ud12 homo sapien
24	37	9.5	41	11	Q99PH5	Q99ph5 mus musculu
25	37	9.5	43	5	Q9W4K6	Q9w4k6 drosophila
26	36.5	9.3	32	4	Q8WTO0	Q8wtq0 homo sapien
27	36.5	9.3	33	4	Q9UIH2	Q9uih2 homo sapien
28	36.5	9.3	37	16	Q927N0	Q927n0 listeria mo
29	36.5	9.3	43	2	O05188	O05188 bacillus su
30	36	9.2	40	12	O91K65	O91k65 hepatitis c
31	36	9.2	42	10	O23471	O23471 arabidopsis
32	36	9.2	43	5	Q95P80	Q95p80 carcinus ma
33	35.5	9.1	32	6	O77493	O77493 lemur catta
34	35.5	9.1	38	4	Q9UGU2	Q9ugu2 homo sapien
35	35.5	9.1	39	5	Q9TX98	Q9tx98 caenorhabdi
36	35.5	9.1	40	12	O57150	O57150 human herpe
37	35.5	9.1	41	12	O81248	O81248 hepatitis c
38	35.5	9.1	41	12	O81252	O81252 hepatitis c
39	35.5	9.1	41	12	O81253	O81253 hepatitis c
40	35.5	9.1	42	3	Q9Y733	Q9y733 candida tro
41	35.5	9.1	42	3	Q9Y734	Q9y734 candida tro
42	35.5	9.1	43	10	Q9LQI3	Q9lqi3 arabidopsis
43	35.5	9.1	43	11	O8R3Y9	O8r3y9 mus musculu
44	35.5	9.1	43	12	O8S597	O8s597 reovirus sp
45	35	9.0	32	10	O8S527	O8s527 ipomoea bat
46	35	9.0	34	4	O14253	O14253 homo sapien
47	35	9.0	36	16	O9KLF9	O9klr9 vibrio chol
48	35	9.0	36	16	Q99QX1	Q99qx1 staphylococ
49	35	9.0	40	12	O8QUT9	O8qut9 infectious
50	35	9.0	42	6	O18958	O18958 bos taurus

## ALIGNMENTS

RESULT 1	Q98TP9	PRELIMINARY:	PRT:	40 AA.
ID	Q98TP9;			
AC	Q98TP9;			
DT	01-JUN-2001 (Tremblrel. 17, Created)			
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)			
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)			
DE	Metallothionein (MT) (Fragment).			
GN	MT.			
OS	Platichthys flesus (European flounder).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;			
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;			
OC	Pleuronectoidei; Pleuronectidae; Platichthys.			
OX	NCBI_TaxID=8260;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=OVARY;			
RA	Williams T.D., Chipman J.K.;			
RT	"A DNA array to monitor the effects of environmental pollution on			
RT	European flounder (Platichthys flesus).";			
RL	Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: METALLOTHIONEINS HAVE A HIGH CONTENT OF CYSTEINE			
CC	RESIDUES THAT BIND VARIOUS HEAVY METALS (BY SIMILARITY).			
CC	-!- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 1.			
DR	EMBL; AJ291833; CAC28138.1; -			
DR	HSSP; P02795; IMHU.			
DR	InterPro; IPR003019; Metallthion.			
DR	InterPro; IPR000006; Metallthion_vert.			
DR	Pfam; PF00131; metalthio; 1.			
DR	PROSITE; PS00203; METALLOTHIONEIN_VRT; 1.			
KW	Chelation; Metal-binding; Metal-thiolate cluster.			
FT	NON_TER 1			
FT	NON_TER 40			

Query Match 13.6%; Score 53; DB 13; Length 39;  
Best Local Similarity 42.9%; Pred. No. 7.7;  
Matches 12; Conservative 3; Mismatches 9; Indels 4; Gaps 2;

QY 21 CDGRCQRCFCNVVVKPL--CIKICAPGC 46  
DB 5 CABQCNRRCRG--PKPIDCCNEHCAGGC 30

RESULT 4  
Q9UDP7 PRELIMINARY; PRT; 41 AA.  
AC Q9UDP7;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE WGSCH\_H\_DJ1062J16.1 protein (Fragment).  
DE WGSCH\_H\_DJ1062J16.1.  
GN Homo sapiens (Human).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9063792; PubMed=9847074;  
RA Sulston J.E., Waterston R.;  
RT "Toward a complete human genome sequence."  
RL Genome Res. 8:1097-1108(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Ali J., Wohlmann P., Ames M., Duckels G.;  
RT "The sequence of Homo sapiens PAC clone RP5-1062J16."  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Waterston R.;  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC006353; AAD28059.1;  
FT NON\_TER 41 41  
SQ SEQUENCE 41 AA; 5038 MW; 69855DB76787619A CRC64;

Query Match 12.4%; Score 48.5; DB 4; Length 41;  
Best Local Similarity 42.3%; Pred. No. 32;  
Matches 11; Conservative 5; Mismatches 3; Indels 7; Gaps 2;

QY 38 CIKICAPGCVCRL----GYLRNKKV 59  
DB 10 CIQLL---CVCRLDWANGYRQRRKL 32

RESULT 5  
Q9QX87 PRELIMINARY; PRT; 37 AA.  
AC Q9QX87;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE MAP19 protein (Fragment).  
GN MAP19.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=FISHER 344;  
RX MEDLINE=20054576; PubMed=10586086;  
RA Stover C.M., Thiel S., Lynch N.J., Schwaebler W.J.;  
RT "The rat and mouse homologues of MASP-2 and MAP19, components of the  
mannan-binding lectin activation pathway of complement."  
RL J. Immunol. 163:6848-6859(1999).  
FT NON\_TER 39 39  
SQ SEQUENCE 39 AA; 4195 MW; 42303BA05F6DB557 CRC64;

QY 7 GGRGKPSNFELSKDRCRCFCNVVVKPLCIKICAPGCVCVR 49  
DB 5 GGSCTC-KNCSCTTCNKSCPCPGCPK-----CASGCVCK 40

RESULT 2  
Q99KF9 PRELIMINARY; PRT; 44 AA.  
AC Q99KF9;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 21, Last annotation update)  
DE Hypothetical 4.9 kDa protein.  
GN 311001K13RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX Strausberg R.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC004681; AA004681.1;  
DR MGD; MGI:1919666; 3110001K13RIK.  
KW Hypothetical protein.  
SQ SEQUENCE 44 AA; 4900 MW; C49523A82D448591 CRC64;

Query Match 16.0%; Score 62.5; DB 11; Length 44;  
Best Local Similarity 33.3%; Pred. No. 0.47;  
Matches 14; Conservative 5; Mismatches 9; Indels 9; Gaps 2;

QY 25 CQFCPCNVVVKPLCIKICAPGCVCRLGYLRNKKVVCVPRSKC 66  
DB 10 CLCVCVSV-----CVSVCSVCVCLCVYMR----AHMPMSKC 42

RESULT 3  
Q9PVG7 PRELIMINARY; PRT; 39 AA.  
AC Q9PVG7;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Receptor tyrosine kinase Xmrk (Fragment).  
GN XMRK.  
OS Xiphophorus maculatus (Southern platyfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.  
OX NCBI\_TaxID=8083;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=SR;  
RX MEDLINE=99126443; PubMed=9927468;  
RA Gutbrod H., Scharl M.;  
RT "Intragenic sex-chromosomal crossovers of Xmrk oncogene alleles affect  
pigment pattern formation and the severity of melanoma in  
Xiphophorus."  
RL Genetics 151:773-783(1999).  
DR EMBL; AF092693; AAD03714.1;  
DR InterPro: IPR002174; Furin-like.  
DR Pfam: PF00757; Furin-like; I.  
KW Kinase.  
FT NON\_TER 1 1  
FT NON\_TER 39 39  
SQ SEQUENCE 39 AA; 4195 MW; 42303BA05F6DB557 CRC64;



```

RESULT 9
Q8WV73 PRELIMINARY; PRT; 40 AA.
AC Q8WV73;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hypothetical 4.5 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC018658; AAI18658.1;
KW Hypothetical protein.
SQ SEQUENCE 40 AA; 4514 MW; 0D53A1D34861EC14 CRC64;

Query Match 11.3%; Score 44; DB 4; Length 40;
Best Local Similarity 36.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 37 LCIKICAPGCVCRGLYLRN 55
Db 3 VCVVCVVCVVCVVCVCIWLK 21

RESULT 10
Q9LW92 PRELIMINARY; PRT; 45 AA.
AC Q9LW92;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Proteinase inhibitor II (Fragment).
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RX MEDLINE=20399434; PubMed=10945337;
RA Hara K., Yagi M., Koizumi N., Kusano T., Sano H.;
RT "Screening of wound-responsive genes identifies an immediate-early
RT expressed gene encoding a highly charged protein in mechanically
RT wounded tobacco plants."
RL Plant Cell Physiol. 41:684-691(2000).
DR EMBL; AB009886; BAA95792.1;
DR InterPro; IPR003465; Prot_inhib.
DR Pfam; PF02428; Prot_inhib-II; 1.
FT NON_TER 1
SQ SEQUENCE 45 AA; 4856 MW; C07AAC196811B7C3 CRC64;

Query Match 11.3%; Score 44; DB 10; Length 45;
Best Local Similarity 36.0%; Pred. No. 1.4e+02;
Matches 9; Conservative 5; Mismatches 7; Indels 4; Gaps 2;

QY 29 CPNV---VPKPLCIKICAPGCVCR 50
Db 10 CPSTQKITKP-CVNCCSPKKGCKL 33

RESULT 11
O18625 PRELIMINARY; PRT; 42 AA.
AC O18625;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)

```

\_\_\_\_\_



RA Zhong X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003615; AAF52441.1; -  
 DR FlyBase: FBgn0042189; CG17376.  
 SQ SEQUENCE 45 AA; 4869 MW; F8F92B3BFB8254 CRC64;

Query Match 10.2%; Score 40; DB 5; Length 45;  
 Best Local Similarity 25.0%; Pred. No. 4.8e+02;  
 Matches 10; Conservative 5; Mismatches 19; Indels 6; Gaps 1;

QY 5 GLGGRKCPSPNEIFSRCD-----GRCQFCFNPVVVVKPLC 38  
 | | | | | : : : : : | : | | | |  
 Db 6 GCCGSPCPRRYLVNKDNAPCVWCAKRAHCYNTPPKCC 45

## RESULT 19

Q94214 ID Q94214 PRELIMINARY; PRT; 35 AA.

AC Q94214; DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE P0455H0310.11 protein.  
 GN P0455H0310.11.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 clone:P0455H0310.";  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP003435; BAB68022.1; -  
 SQ SEQUENCE 35 AA; 3068 MW; C6F1E2AB366A6B51 CRC64;

Query Match 10.0%; Score 39; DB 10; Length 35;  
 Best Local Similarity 58.3%; Pred. No. 5.2e+02;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 GFGGLGGRKGC 12  
 | | | | |  
 Db 24 GGAAGCGSGGC 35

## RESULT 20

Q8XTX4 ID Q8XTX4 PRELIMINARY; PRT; 42 AA.

AC Q8XTX4; DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Hypothetical protein RSC3422.  
 GN RSC3422 OR RS01802.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
 OC Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GM11000;  
 RX MEDLINE=21681879; PubMed=11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,  
 RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,  
 RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,  
 RA Weissenbach J., Boucher C.A.;

RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*."  
 RL Nature 415:497-502(2002).  
 DR EMBL: AL646075; CAD16919.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 42 AA; 4646 MW; FD0483199B20FD51 CRC64;

Query Match 10.0%; Score 39; DB 16; Length 42;  
 Best Local Similarity 38.5%; Pred. No. 6.1e+02;  
 Matches 5; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 17 IFSRCDGRCORFC 29  
 : : | | | : : : |  
 Db 8 LWSGCDGKAQSC 20

## RESULT 21

Q8W2G7 ID Q8W2G7 PRELIMINARY; PRT; 38 AA.

AC Q8W2G7; DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE 4-hydroxyphenylpyruvate dioxygenase (Fragment).  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. TAINONG 67; TISSUE=LEAF;  
 RX MEDLINE=21326245; PubMed=11432928;  
 RA Lee R.H., Wang C.H., Huang L.T., Chen S.C.;  
 RT "Leaf senescence in rice plants: cloning and characterization of  
 senescence up-regulated genes.";  
 RL J. Exp. Bot. 52:1117-1121(2001).  
 DR EMBL: AF251065; AAL65390.1; -  
 KW Dioxygenase; Pyruvate.  
 FT NON\_TER  
 SQ SEQUENCE 38 AA; 4037 MW; 627174633993A0FE CRC64;

Query Match 9.7%; Score 38; DB 10; Length 38;  
 Best Local Similarity 53.3%; Pred. No. 7.6e+02;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 GGLGGRKCPSPNEIF 18  
 | | | | | : | |  
 Db 3 GCGGFGKGNFSELF 17

## RESULT 22

Q9MRL7 ID Q9MRL7 PRELIMINARY; PRT; 42 AA.

AC Q9MRL7; DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Putative ribulose-phosphate 3-epimerase (Fragment).  
 GN RPE.  
 OS Daucus carota (Carrot).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.  
 OX NCBI\_TaxID=4039;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. ST. VALERY; TISSUE=CALLUS;  
 RA Boschi E., Filippini F., Lo Schiavo F., Terzi M., Vergara R.;  
 RT "Identification of differential transcripts in auxinR lines of *Daucus  
 carota* L.";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ297425; CAB97126.1; -

DR HSSP: Q43843: 1RPX.  
 DR InterPro: IPR000056; Ribul\_P\_3\_epim.  
 DR Pfam: PF00834; Ribul\_P\_3\_epim; 1.  
 DR PROSITE: PS01085; RIBUL\_P\_3\_EPIMER\_1; 1.  
 KW Chloroplast.  
 FT NON\_TER 1 1  
 FT 42 42  
 SQ SEQUENCE 42 AA: 4579 MW; 0EA026C8F51E14F3 CRC64;  
 Query Match 9.6%; Score 37.5; DB 8; Length 42;  
 Best Local Similarity 39.1%; Pred. No. 9.7e+02;  
 Matches 9; Conservative 2; Mismatches 9; Indels 3; Gaps 1;  
 QY 22 DGRCORFCNVVPKPLCIKICAP 44  
 DB 12 DG---RFPVNITIGPLVDALRP 31  
 RESULT 23  
 Q3UD12 PRELIMINARY; PRT: 33 AA.  
 AC Q3UD12; 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE AMGX protein (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95322983; PubMed=7599636;  
 RA Lench N.J., Winter G.B.;  
 RT "Characterisation of molecular defects in X-linked amelogenesis imperfecta (ATHL)."  
 RT Hum. Mutat. 5:251-259(1995).  
 RL InterPro: IPR001230; Pfam: Ribul\_P\_3\_epim.  
 DR PROSITE: PS00294; PRNLYATION; UNKNOWN\_1.  
 SQ SEQUENCE 33 AA: 3681 MW; D131F784BD7D8C93 CRC64;  
 Query Match 9.5%; Score 37; DB 4; Length 33;  
 Best Local Similarity 30.8%; Pred. No. 9e+02;  
 Matches 8; Conservative 3; Mismatches 15; Indels 0; Gaps 0;  
 QY 41 ICAPGCVCRGLYLRNKKKVCVPRSK 66  
 DB 5 LCTPCSPCRHSHLCRCSPCLPC 30  
 RESULT 24  
 Q99PH5 PRELIMINARY; PRT: 41 AA.  
 AC Q99PH5; 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE Insulinoma associated-2 (Fragment).  
 GN PTPN OR IA-2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SVJ;  
 RX MEDLINE=21012487; PubMed=11126414;  
 RA Saeki K., Xie J., Notkins A.L.;  
 RT "Genomic structure of mouse IA-2: comparison with its human homologue."  
 RT Diabetologia 43:1429-1434(2000).  
 RL EMBL: AF288816; AAK07090.1;  
 DR MGD: MGI:102765; Ptpn.

FT NON\_TER 41  
 SQ SEQUENCE 41 AA: 4018 MW; BDD29522E5247C43 CRC64;  
 Query Match 9.5%; Score 37; DB 11; Length 41;  
 Best Local Similarity 70.0%; Pred. No. 1.1e+03;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 GFGFGLGGRG 10  
 DB 8 GSGSGSGSG 17  
 RESULT 25  
 Q9W4K6 PRELIMINARY; PRT: 43 AA.  
 AC Q9W4K6; 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)  
 DE CG12687 protein.  
 GN CG12687.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton R.G., Mortman J.R., Fendell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,  
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M., Reese M.G.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Smith H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003431; AAF45945.1;  
 DR FlyBase: FBgn0040908; CG12687.  
 SQ SEQUENCE 43 AA: 4461 MW; 4A22167D4CFBAE45 CRC64;  
 Query Match 9.5%; Score 37; DB 5; Length 43;



Thu Feb 27 09:12:58 2003

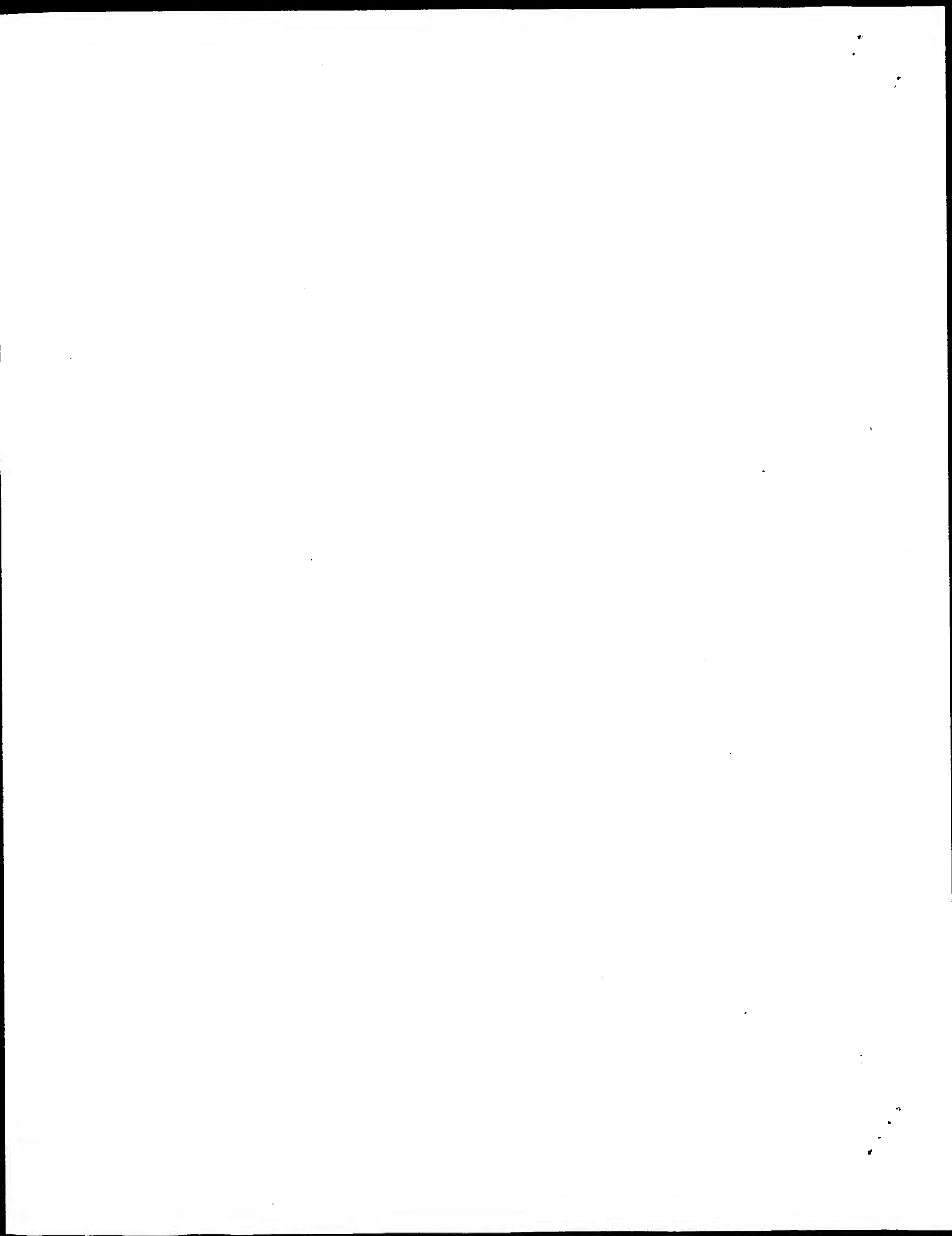
us-09-506-978-1.slim.rspt

Page 9

Best Local Similarity 45.5%; Pred. No. 1.2e+03;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 24 RCORFCPNVVP 34  
Db :| ||| :|  
3 KCLRFCVGTIP 13

Search completed: February 27, 2003, 09:03:44  
Job time : 32 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

COM protein - protein search, using sw model

```
Run on: February 27, 2003, 09:03:04 ; Search time 33 Seconds
        (without alignments)
        270.539 Million cell updates/sec
```

Title: US-09-506-978-1

Perfect score:

Sequence: 1 GGFGGLGGRGKCPSNEIFSR.....CRLGYLRNKKKVCVPRSKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 63089

Minimum DB seq length: 32

Maximum DB seq length: 45

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

```

Database :
1: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	62	15.9	43	22	AAB60286	Human factor IX (h
2	58.5	15.0	39	20	AAV02082	KIX peptide used t
3	57	14.6	40	5	AAP40692	Sequence encoded b
4	57	14.6	40	5	AAP40220	Partial sequence o
5	53	13.6	38	22	ABB95640	Human testicular a
6	53	13.6	38	22	AAAG4179	Human reproductive
7	52.5	13.4	36	23	AAU91021	Transplant media a
8	51	13.0	32	17	AAW12903	Modified ligand sp
9	50	12.8	32	21	AAV59107	N. clavipes spider
10	50	12.8	38	20	AAV02077	Kalio toxin peptide

11	49	12.5	38	20	AAV02081	KTX peptide used
12	49	12.5	40	23	AAU91037	Transplant media a
13	49	12.5	42	19	AAW48751	T-cell surface ant
14	49	12.5	41	21	AAV88018	Human growth hormo
15	48.5	12.4	34	22	AAO09332	Human polypeptide
16	48.5	12.4	9	22	AAW82390	KM-2 antiviral and
17	48.5	12.4	44	22	AAO09473	Human polypeptide
18	48	12.3	37	22	AAW92225	Toxin peptide SEQ
19	48	12.3	37	22	AAW92226	Toxin peptide SEQ
20	48	12.3	40	15	AAW63519	Bovine neutrophil
21	48	12.3	40	23	AAU91032	Transplant media a
22	48	12.3	41	22	AAW84582	Amino acid sequenc
23	48	12.3	42	22	AAW84587	Amino acid sequenc
24	47	12.0	39	18	AAV31658	Nisin-subtilin chl
25	47	12.0	39	22	ABG23312	Novel human diagno
26	47	12.0	45	18	AAW01613	Alpha-hordothionin
27	47	12.0	45	19	AAW41690	Methionine substit
28	46.5	11.9	38	20	AAV02080	Agitoxin peptide u
29	46.5	11.9	38	21	AAV83190	IRK1 Agitoxin2 lig
30	46.5	11.9	41	22	ABW40595	Peptide #8101 enco
31	46.5	11.9	41	22	AAW61453	Human brain expres
32	46.5	11.9	41	22	AAW74243	Human bone marrow
33	46.5	11.9	41	22	AAW34355	Peptide #8392 enco
34	46	11.8	32	19	AAW375704	Metal binding prot
35	46	11.8	33	19	AAW53350	Nephila clavipes s
36	46	11.8	33	21	AAV59074	N. clavipes spider
37	46	11.8	42	22	AAO13358	Human polypeptide
38	46	11.8	43	22	ABG28039	Novel human diagno
39	46	11.8	43	22	AAU14870	Novel bone marrow
40	46	11.8	45	14	AAW42283	Corticostatin. Sy
41	45.5	11.6	40	22	AAO02434	Human polypeptide
42	45.5	11.6	41	20	AAV27030	Amino acid sequenc
43	45	11.5	32	17	AAW12904	Modified ligand sp
44	45	11.5	33	22	AAW91722	Human immune/haema
45	45	11.5	39	18	AAW24383	Gallinacin 1. Gal
46	45	11.5	45	15	AAW60047	Alpha-hordothionin
47	45	11.5	45	15	AAW60051	Alpha-hordothionin
48	45	11.5	45	15	AAW60052	Alpha-hordothionin
49	44.5	11.5	45	19	AAW58555	High lysine deriva
50	44.5	11.4	45	22	AAO05239	Human polypeptide

## ALIGNMENTS

RESULT 1  
AAB60286

ID AAB60286 standard; Protein: 43 AA.

AA  
AC AAB60286.

30-MAR-2003 16:44

XX

XXXXXXXXXX (MTR) CASH ENCODED PAYMENT, SEQ ID NO:3.  
XXXX

Age-related gene regulation; liver-specific; gene expression;

human factor IX; hFIX; AE5'; AE3'; age-regulatable expression construct;

antisense therapy; gene therapy; thrombosis; cardiovascular disease;

diabetes; Alzheimer's disease; parkinson's disease; cancer; osteoporosis;

XXXX

XX  
nono sapientis.

WO2000/5279-

14-DEC-2000.

06-JUN-2000; 2000WO-US15728.

09-JUN-1999 0905-03

XX 7A

[illegible]

```

XX PD 16-FEB-1984.
XX PF 03-AUG-1983; 83WO-GB00191.
XX PR 06-MAY-1983; 83GB-0012491.
XX PR 04-AUG-1982; 82GB-0022485.
XX PA (NATR ) NATIONAL RES DEV CORP.
XX PA (BROW/) BROWNLEE G G.
XX PT Brownlee G, Choo KH;
XX DR WPI; 1984-049331/08.
XX DR N-PSDB; AAN40141.
XX
XX PT Recombinant DNA cloning vehicles - useful in prodn. of factor IX
XX PT polypeptide and of diagnostic probes for Christmas disease
XX PS Example; Fig 7a-i; 7lpp; English.
XX
XX CC The inventors claim DNA molecules comprising part or all of the
XX CC human factor IX DNA. The invention also includes cDNA derived from
XX CC human factor IX RNA. Specifically claimed are: recombinant DNA (the
XX CC phage present in clone lambda HIX-1) deposited as NCIB No. 11749;
XX CC Recombinant DNA in which the cloning vehicle is the modified PAT 153
XX CC plasmid present in E.coli NCIB No. 11747; Recombinant DNA in which
XX CC the bovine factor IX DNA sequence is contained in the recombinant
XX CC DNA transformed into E.coli to form a clone deposited as NCIB No.
XX CC 11748.
XX
XX SQ Sequence 40 AA;

Query Match 14.6%; Score 57; DB 5; Length 40;
Best Local Similarity 32.5%; Pred. No. 78;
Matches 13; Conservative 5; Mismatches 14; Indels 8; Gaps 2;

QY 22 DGRCORFCPNVVPKPLCIKICAPGCVCRIGY-LRNKKKVC 60
Db :|||:| | | | | | | | | | | | | | | |
8 NGRCEQFCNKSADNKVV-----CSTEGYRLAENQKSC 40

RESULT 4
AAP40220
ID AAP40220 standard; Protein; 40 AA.
AC AAP40220;
XX
XX DT 13-FEB-1992 (first entry)
XX
XX DE Partial sequence of human factor IX (starting AA no. 85) encoded
XX DE by an exon of human factor IX genomic DNA.
XX
XX KW Haemophilia; Christmas disease; diagnosis; treatment.
XX
XX OS Homo sapiens.
XX
XX PN GB2125409-A.
XX
XX PD 07-MAR-1984.
XX
XX PF 03-AUG-1983; 83GB-0020975.
XX
XX PR 16-MAY-1983; 83GB-0012490.
XX PR 04-AUG-1982; 82GB-0022486.
XX PR 03-AUG-1983; 83GB-0020975.
XX
XX PA (NATR ) NATIONAL RES DEV CORP.
XX
XX PI Brownlee GG, Choo KH;
XX
XX DR WPI; 1984-057898/10.
XX DR N-PSDB; AAN40176.

```

```

XX
XX PT Prodn. of artificial human factor IX - by use of recombinant DNA
XX PT sequences for host transformation and cultivation
XX
XX PS Example; Fig 7; 49pp; English.
XX
XX CC The inventors claim a recombinant DNA having a human factor IX
XX CC sequence pref. at least 50 nucleotides long, esp. 75-27000
XX CC nucleotides. A cloning vector contg. foreign DNA is also claimed.
XX CC The foreign sequence pref. includes the whole of an exon sequence of
XX CC the human factor IX genome. The cloning vehicle may be a modified
XX CC PAT 153 plasmid. Also claimed is a labelled diagnostic probe
XX CC comprising a DNA molecule having a single- or double-stranded probe
XX CC sequence of 45 to 10000 nucleotides long Factor IX DNA sequence.
XX
XX SQ Sequence 40 AA;

Query Match 14.6%; Score 57; DB 5; Length 40;
Best Local Similarity 32.5%; Pred. No. 78;
Matches 13; Conservative 5; Mismatches 14; Indels 8; Gaps 2;

QY 22 DGRCORFCPNVVPKPLCIKICAPGCVCRIGY-LRNKKKVC 60
Db :|||:| | | | | | | | | | | | | | | |
8 NGRCEQFCNKSADNKVV-----CSTEGYRLAENQKSC 40

RESULT 5
ABB95640
ID ABB95640 standard; Protein; 38 AA.
XX
XX AC ABB95640;
XX
XX DT 21-JUN-2002 (first entry)
XX
XX DE Human testicular antigen SEQ ID NO: 1024.
XX
XX KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
XX KW reproductive system disorder; urinary system disorder; gene therapy;
XX KW cardiovascular disorder; respiratory disorder; neurological disorder;
XX KW gastrointestinal disease; infection; cytostatic.
XX
XX OS Homo sapiens.
XX
XX PN WO200155317-A2.
XX
XX PD 02-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US01329.
XX
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 11-JUL-2000; 2000US-0217496.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.

```

PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232337.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-483232/52.  
XX Nucleic acids encoding 973 human testicular antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating testicular cancer  
XX Claim 11; SEQ ID NO 1024; 766pp; English.  
XX The present invention provides the protein and coding sequences of 973  
XX human testicular antigens, and fragments of their genomic sequences. The  
XX sequences can be used in the treatment of cardiovascular, urinary system,  
XX reproductive system, immune, respiratory, neurological and  
XX gastrointestinal disorders, infections, and particularly cancer,  
XX especially testicular cancers. The present sequence is a protein of the  
XX invention.

XX Sequence 38 AA;  
XX Query Match 13.6%; Score 53; DB 22; Length 38;  
XX Best Local Similarity 35.6%; Pred. No. 1.9e+02;  
XX Matches 16; Conservative 5; Mismatches 14; Indels 10; Gaps 3;  
Qy 6 LGGRGKCPSEIFSRGRCORCFPNVVPKLCIKICAPGVCRL 50  
| | | | | | | | | | | | | | | | | | | | | |  
Db 2 LDARASWPSNQL---CDLR-----QVIP-ALCEFLICSMGMIIQL 36

RESULT 6  
AAM94179  
ID AAM94179 standard; Protein; 38 AA.

XX	AC	AA941179;	PR	14-SEP-2000;	2000US-0232400
XX	DT	21-NOV-2001 (first entry)	PR	14-SEP-2000;	2000US-0232401
XX	DE	Human reproductive system related antigen SEQ ID NO: 2837.	PR	14-SEP-2000;	2000US-0233063
XX	KW	Human; reproductive system related antigen; reproductive system disorder;	PR	14-SEP-2000;	2000US-0233064
XX	KW	cancer; gene therapy.	PR	14-SEP-2000;	2000US-0233065
XX	OS	Homo sapiens.	PR	21-SEP-2000;	2000US-0234223
XX	PN	WO200155320-A2.	PR	21-SEP-2000;	2000US-0234297
XX	PD	02-AUG-2001.	PR	25-SEP-2000;	2000US-0234958
XX	PF	17-JAN-2001; 2001WO-US01339.	PR	26-SEP-2000;	2000US-0234984
XX	PR	31-JAN-2000; 2000US-0179065.	PR	27-SEP-2000;	2000US-0235834
PR	PR	04-FEB-2000; 2000US-0180628.	PR	29-SEP-2000;	2000US-0235836
PR	PR	24-FEB-2000; 2000US-0184664.	PR	29-SEP-2000;	2000US-0236327
PR	PR	02-MAR-2000; 2000US-0186350.	PR	29-SEP-2000;	2000US-0236367
PR	PR	16-MAR-2000; 2000US-0189874.	PR	29-SEP-2000;	2000US-0236368
PR	PR	17-MAR-2000; 2000US-0190076.	PR	29-SEP-2000;	2000US-0236369
PR	PR	18-APR-2000; 2000US-0198123.	PR	29-SEP-2000;	2000US-0236370
PR	PR	19-MAY-2000; 2000US-0205515.	PR	02-OCT-2000;	2000US-0236802
PR	PR	07-JUN-2000; 2000US-0209467.	PR	02-OCT-2000;	2000US-0237037
PR	PR	28-JUN-2000; 2000US-0214886.	PR	02-OCT-2000;	2000US-0237038
PR	PR	30-JUN-2000; 2000US-0215135.	PR	02-OCT-2000;	2000US-0237039
PR	PR	07-JUL-2000; 2000US-0216647.	PR	13-OCT-2000;	2000US-0237040
PR	PR	07-JUL-2000; 2000US-0216860.	PR	13-OCT-2000;	2000US-0239935
PR	PR	11-JUL-2000; 2000US-0217487.	PR	20-OCT-2000;	2000US-0239937
PR	PR	11-JUL-2000; 2000US-0217496.	PR	20-OCT-2000;	2000US-0240960
PR	PR	14-JUL-2000; 2000US-0218290.	PR	20-OCT-2000;	2000US-0241221
PR	PR	26-JUL-2000; 2000US-0220963.	PR	20-OCT-2000;	2000US-0241785
PR	PR	14-AUG-2000; 2000US-0224518.	PR	20-OCT-2000;	2000US-0241786
PR	PR	14-AUG-2000; 2000US-0224519.	PR	20-OCT-2000;	2000US-0241787
PR	PR	14-AUG-2000; 2000US-0225213.	PR	20-OCT-2000;	2000US-0241808
PR	PR	14-AUG-2000; 2000US-0225266.	PR	20-OCT-2000;	2000US-0241809
PR	PR	14-AUG-2000; 2000US-0225267.	PR	20-OCT-2000;	2000US-0241826
PR	PR	14-AUG-2000; 2000US-0225268.	PR	01-NOV-2000;	2000US-024617
PR	PR	14-AUG-2000; 2000US-0225270.	PR	08-NOV-2000;	2000US-0246474
PR	PR	14-AUG-2000; 2000US-0225447.	PR	08-NOV-2000;	2000US-0246475
PR	PR	14-AUG-2000; 2000US-0225757.	PR	08-NOV-2000;	2000US-0246476
PR	PR	14-AUG-2000; 2000US-0225758.	PR	08-NOV-2000;	2000US-0246477
PR	PR	18-AUG-2000; 2000US-0225759.	PR	08-NOV-2000;	2000US-0246478
PR	PR	22-AUG-2000; 2000US-0226279.	PR	08-NOV-2000;	2000US-0246523
PR	PR	22-AUG-2000; 2000US-0226681.	PR	08-NOV-2000;	2000US-0246524
PR	PR	22-AUG-2000; 2000US-0226681.	PR	08-NOV-2000;	2000US-0246525
PR	PR	22-AUG-2000; 2000US-0226868.	PR	08-NOV-2000;	2000US-0246526
PR	PR	22-AUG-2000; 2000US-0227182.	PR	08-NOV-2000;	2000US-0246527
PR	PR	23-AUG-2000; 2000US-0227009.	PR	08-NOV-2000;	2000US-0246528
PR	PR	30-AUG-2000; 2000US-0228924.	PR	08-NOV-2000;	2000US-0246532
PR	PR	01-SEP-2000; 2000US-0229287.	PR	08-NOV-2000;	2000US-0246610
PR	PR	01-SEP-2000; 2000US-0229343.	PR	08-NOV-2000;	2000US-0246611
PR	PR	01-SEP-2000; 2000US-0229344.	PR	17-NOV-2000;	2000US-0249207
PR	PR	05-SEP-2000; 2000US-0229345.	PR	17-NOV-2000;	2000US-0249208
PR	PR	05-SEP-2000; 2000US-0229509.	PR	17-NOV-2000;	2000US-0249209
PR	PR	05-SEP-2000; 2000US-0229513.	PR	17-NOV-2000;	2000US-0249210
PR	PR	06-SEP-2000; 2000US-0230437.	PR</		





XX Harkins RN, Katz BA, Seto M;  
 XX WPI; 1996-371432/37.  
 DR Modified ligand specific for Class I receptor tyrosine kinase  
 XX used to quantify the presence of the kinase on human cells, and to  
 XX treat malignancy and aberrant expression  
 PS Claim 13; Page 44; 58pp; English.  
 XX Modified ligand specific for a Class I receptor tyrosine kinase (RTK)  
 CC and corresponding to a native RTK ligand with amino acids forming 3 Cys  
 CC linkages is claimed. In the native RTK ligand these linkages define  
 CC domains A, B and C resp. and have 6 Cys residues, with the first linkage  
 CC being between Cys1 and Cys3, the second between Cys2 and Cys4, and  
 CC the third between Cys5 and Cys6. Domain A is Cys1, Cys2 and the amino  
 CC acids between; domain B is the amino acids between Cys2 and Cys4; and  
 CC domain C is Cys4 and the amino acids attached to the C-terminal of Cys4.  
 CC The new modified RTK ligand has 4 Cys residues forming 2 linkages  
 CC between Cys2 and 4 and Cys5 and 6 resp., and has 2 domains, A1 and C1,  
 CC corresponding to domains A and C resp. of the native ligand and contg.  
 CC the biologically active amino acids of these domains. Cys 2, which is  
 CC the C-terminus of domain A1, and Cys 4, which is the N-terminus of C1,  
 CC are connected by a peptide bridge comprising a prebridge subdomain and  
 CC a bridge subdomain. Within the peptide bridge there is a beta turn  
 CC corresponding to that present within domain B of the native RTK ligand.  
 CC The modified ligand is used to quantify the presence of RTKs on a human  
 CC cell, and to treat malignancy and aberrant expression. It may have  
 CC agonist or antagonist activity. The modified ligand can be labelled  
 CC easily to provide a rapid assay for the presence or absence of RTKs on  
 CC particular cells.  
 CC The present sequence is one of 10 specific examples of the new ligands  
 CC directed towards the EGF receptor.  
 XX Sequence 32 AA;  
 SQ

Query Match 13.0%; Score 51; DB 17; Length 32;  
 Best Local Similarity 50.0%; Pred. No. 2.5e+02;  
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 38 CIKICAPGCVCRGLYL 53  
 I :||||| I:  
 Db 6 CPHGTSPGCVCHSGYV 21

RESULT 9  
 AAY59107  
 ID AAY59107 standard; peptide; 32 AA.  
 XX  
 AC AAY59107;  
 XX  
 DT 08-MAR-2000 (first entry)  
 XX  
 DE N. clavipes spider silk protein 1 repeat fragment.  
 XX  
 KW Spider silk protein; dragline silk protein; fiber.  
 XX  
 OS Nephila clavipes.  
 XX  
 PN US5989894-A.  
 XX  
 PD 23-NOV-1999.  
 XX  
 PF 04-OCT-1994; 94US-0317844.  
 XX  
 PR 15-APR-1991; 91US-0684819.  
 XX  
 PR 20-APR-1990; 90US-0511792.  
 XX  
 PA (UYWY-) UNIV WYOMING.  
 XX  
 PI Hinman MB, Xu M, Lewis RV;  
 XX

DR WPI; 2000-061225/05.  
 XX Isolated DNA, vector and transformed cell encoding for and useful in  
 XX the production of spider silk protein  
 XX  
 PS Claim 5; Columns 71-72; 65pp; English.  
 XX  
 CC The invention provides isolated cDNA molecules coding for spider silk  
 CC proteins. The spider silk proteins are characterized by repeating alpha  
 CC and beta regions and optional variable regions. The DNA sequences are  
 CC useful in the production of spider silk protein by recombinant DNA  
 CC techniques. The recombinant spider silk proteins may be used for the  
 CC production of fibers. Sequences AAY59101-125 represent repeat fragments  
 CC of N. clavipes spider silk protein 1.  
 XX  
 SQ Sequence 32 AA;  
 Query Match 12.8%; Score 50; DB 21; Length 32;  
 Best Local Similarity 80.0%; Pred. No. 3.2e+02;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGGLGGRG 10  
 II:||||| I:  
 Db 23 GGYGLGGGG 32

RESULT 10  
 AAY02077  
 ID AAY02077 standard; peptide; 38 AA.  
 XX  
 AC AAY02077;  
 XX  
 DT 06-JUL-1999 (first entry)  
 XX  
 DE Kallitoxin peptide used to block potassium channels.  
 XX  
 KW Neuronal potassium channel; Kv1.1; Kv1.2; Kv1.3; neurological disorder;  
 KW immune disorder; nerve signal transmission; block; multiple sclerosis;  
 KW cerebral; medullary inflammatory disorder; Baló; Devic; disseminated;  
 KW haemorrhagic acute encephalitis; optic neuropathy; systemic disease;  
 KW Behcet; sarcoidosis; vasculitis; peripheral inflammatory disorder;  
 KW polyradiculoneuritis; infection; collagenosis; neurodegenerative disease;  
 KW amyotrophic lateral sclerosis; metabolic disorder; vascular ischaemia;  
 KW tumour; trauma.  
 XX  
 OS Synthetic.  
 XX  
 PN FR2769226-A1.  
 XX  
 PD 09-APR-1999.  
 XX  
 PF 06-OCT-1997; 97FR-0012421.  
 XX  
 PR 06-OCT-1997; 97FR-0012421.  
 XX  
 PA (CNRS) CNRS CENT NAT RECH SCI.  
 XX  
 PI Beraud E, Crest M, Gola M, Van Rietschoten J;  
 XX  
 DR WPI; 1999-246903/21.  
 XX  
 XX Use of neuronal potassium channel blockers, especially kalliotoxin  
 XX peptides - to prepare therapeutic compositions for treatment of  
 XX neurological disorders  
 PS Claim 8; Page 19; 30pp; French.  
 XX  
 CC The specification describes the use of compounds that block neuronal  
 CC potassium channels of the Kv1.1, Kv1.2 and optionally Kv1.3 type to  
 CC prepare therapeutics for treating human or animal neurological  
 CC disorders, e.g. of immune origin, in which nerve signal transmission  
 CC is blocked. The products are especially used for treating multiple  
 CC sclerosis. Other uses are for treating other cerebral and medullary





KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

XX 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AAI89323.

XX Isolated nucleic acids and polypeptides, useful for preventing  
 PT diagnosing and treating e.g. leukaemia, inflammation and immune  
 PT disorders.

XX Claim 20; SEQ ID NO 23284; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 34 AA;

XX Query Match 12.4%; Score 48.5; DB 22; Length 34;

XX Best Local Similarity 46.2%; Pred. No. 4.7e+02;

XX Matches 12; Conservative 1; Mismatches 2; Indels 11; Gaps 2;

QY 37 LCIR-ICAPGVCVRLGRLNKKKVCV 61

DB 2 LCIRKVCACVC-----VCV 17

RESULT 16

AA082390

ID AAR82390 standard; peptide; 42 AA.

XX AAR82390;

XX 31-OCT-1990 (first entry)

XX KM-2 antiviral and antibacterial peptide.

XX Flesh fly; embryo; antibacterial; antiviral; herpes simplex;  
 KW food additive.

XX Sarcophaga peregrina.

XX EP280859-A.

XX 07-SEP-1987.

XX 20-JAN-1988; 88EP-0100801.

XX 23-JAN-1987; 87JP-0014807.

XX (SANW ) SANWA KAGAKU KENKYUSHO CO LTD.

XX Natori S;

XX WPI; 1988-251489/36.

XX Antiviral polypeptide contg. 40 amino acid residues - obtd. by  
 PT culturing cell line established from Sarcophaga peregrina embryo.

XX Claim 1; Page 3; 8pp; English.

XX The polypeptide is secreted in the body fluid of Sarcophaga peregrina  
 CC when the larva is injured. It has a wide antibacterial spectrum and  
 CC almost no toxicity and excellent thermal stability so may be useful as  
 CC a food additive. It may also be used therapeutically at a dosage of  
 CC 1-500 mg/day.

XX Sequence 42 AA;

XX Query Match 12.4%; Score 48.5; DB 9; Length 42;

XX Best Local Similarity 46.2%; Pred. No. 5.7e+02;

XX Matches 12; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

QY 42 CAPGVCV---RLGRLNKKKVCVPRS 64

DB 17 CAACHLLRNGRGYCNNGRAVCVCRN 42

RESULT 17

AAO09473

ID AAO09473 standard; Protein; 44 AA.

XX AAO09473;

XX 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 23365.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

XX 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AAI89404.

XX Isolated nucleic acids and polypeptides, useful for preventing  
 PT diagnosing and treating e.g. leukaemia, inflammation and immune  
 PT disorders.

XX Claim 20; SEQ ID NO 23365; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and

CC Inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).

2;

## RESULT 19

AA AAB92226:

DE Toxin peptide SEQ ID NO:1402.

DE Toxin peptide SEQ ID NO:1402.

DE Toxin peptide SEQ ID NO:1402.

OS Homo sapiens.

OS Homo sapiens.

AX  
PN  
W0200069900-A2.

XX  
PD  
23-NOV-2000

17-MAY-2000: 2000WQ-UIS13576

XX  
PR 17-MAY-1999. 99JIS-0134406XX  
PR 17-MAY-1999. 99JIS-0134406

PR 10-SEP-1999; 99US-0153406.  
 PR 15-OCT-1999; 99US-0150793

PA (CONJ-) CONJUCHEM INC.

PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibodeau K:

XX  
DR  
WPT: 2001-112059/12

Modifying and attaching therapeutic peptides to albumin prevents  
peptidase degradation, useful for increasing length of *in vivo* activity

PS Disclosure: page 655: 733pp: English:

The present invention describes a modified therapeutic peptide (I) comprising a therapeutically active amino acid region (III) and a reactive group (II) (e.g. succinimidyl and maleimido groups) attached to a less therapeutically active amino acid region (IV), which covalently bonds with amino/hydroxyl/thiol groups on blood components to form a peptidase stabilised therapeutic peptide composed of 3-50 amino acids. (I) are useful for modifying therapeutic peptides e.g. hormones, growth factors and neurotransmitters, to protect them from peptidase activity in vivo for the treatment of various disorders. Endogenous therapeutic peptides are not suitable as drug candidates as they require frequent administration due to rapid degradation by peptidases in the body. Modifying and attaching therapeutic peptides to albumin prevents or reduces the action of peptidases to increase length of activity (half life) and specificity as bonding to large molecules decreases





PD 03-APR-1997.  
 XX  
 PF 30-SEP-1996; 96WO-0515160.  
 XX  
 PR 28-SEP-1995; 95US-0535494.  
 XX  
 PA (UYMA-) UNIV MARYLAND BALTIMORE.  
 XX  
 PI Hansen JN;  
 XX  
 DR WPI; 1997-225847/20.  
 DR N-PSDB; AAX87828.  
 XX  
 PT Lantibiotic mutants and chimera(s) - having enhanced stability and  
 PT activity compared to nisin  
 XX  
 PS Example; Fig 2; 60pp; English.  
 XX  
 CC The present sequence represents a chimeric lantibiotic pre-peptide  
 CC composed of a subtilin leader region and a nisin-subtilin fusion  
 CC comprising residues 1-11 of Lactococcus lactis nisin and residues  
 CC 12-32 of Bacillus subtilis subtilin. DNA (see AAX87828) encoding the  
 CC chimera was produced by mutagenesis of subtilin DNA such that  
 CC residues 1, 2 and 4 of subtilin were replaced by the corresponding  
 CC residues of nisin. The chimera was efficiently processed in B.  
 CC subtilis into a functional lantibiotic. Post-translational  
 CC modifications included the dehydration of serine and threonine  
 CC residues and formation of thioether crosslinkages. The chimera had  
 CC similar activity to nisin with respect to activity against Bacillus  
 CC cereus spores and vegetative cells. The chimera can be produced by  
 CC cultivation of transformed host cells and used e.g. as a food  
 CC preservative to treat, kill or inhibit the growth of microorganisms  
 CC and/or their spores.  
 XX  
 SQ Sequence 39 AA;  
 Query Match 12.0%; Score 47; DB 18; Length 39;  
 Best Local Similarity 43.8%; Pred. No. 7.5e+02;  
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
 OY 32 VVPKPLCIKICAPGCV 47  
 : : : : :  
 Db 4 ITPQITSISLCTPGCV 19  
 RESULT 25  
 ABG23312  
 ID ABG23312 standard; Protein; 39 AA.  
 XX  
 AC ABG23312;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #23303.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US08631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX

DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS87499.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity  
 XX  
 PS Claim 20; SEQ ID No 53671; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG0377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 39 AA;  
 Query Match 12.0%; Score 47; DB 22; Length 39;  
 Best Local Similarity 29.4%; Pred. No. 7.5e+02;  
 Matches 15; Conservative 5; Mismatches 5; Indels 26; Gaps 4;  
 OY 19 SRCD-CRCQRF-----CPNVVVKPLCIKICAPGCVCL-----GYLRN 55  
 : : : : :  
 Db 1 SKCSIGLYORYPYSGNCXN-----SGCICRVYXCCREHRYGVN 39  
 : : : : :  
 Search completed: February 27, 2003, 09:04:23  
 Job time : 35 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 27, 2003, 09:04:28 : Search time 13 seconds  
(without alignments)  
194.412 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GGFGLGGKGCPSNEIFSR.....CRLGVLRNKKVCVPRSKG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174566 segs, 37721826 residues

Total number of hits satisfying chosen parameters: 8992

Minimum DB seq length: 32

Maximum DB seq length: 45

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Database : Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	55.5	14.2	38	10	US-09-847-185-47
2	55.5	14.2	45	9	US-10-136-573A-11
3	55.5	14.2	45	9	US-09-877-665-11
4	55.5	14.2	45	9	US-10-215-862-11
5	55.5	14.2	45	10	US-09-817-647-11
6	49	12.5	40	10	US-09-917-340-73
7	48	12.3	40	10	US-09-917-340-68
8	47.5	12.1	42	10	US-09-865-578-4
9	46.5	11.9	41	10	US-09-864-761-44588
10	46	11.8	37	10	US-09-829-481-11
11	46	11.8	37	10	US-09-917-340-57
12	45	11.5	45	9	US-10-136-573A-13
13	45	11.5	45	9	US-09-877-665-13
14	45	11.5	45	9	US-10-215-862-13
15	45	11.5	45	10	US-09-817-647-13
16	44	11.3	34	10	US-09-040-518-2
17	44	11.3	38	10	US-09-030-619-200
18	44	11.3	40	12	US-10-124-557-8
19	43.5	11.1	37	9	US-09-814-452-5

20	43.5	11.1	45	9	US-10-136-573A-16
21	43.5	11.1	45	9	US-09-877-665-16
22	43.5	11.1	45	9	US-10-215-862-16
23	43.5	11.1	45	10	US-09-817-647-16
24	43	11.0	38	10	US-09-030-619-197
25	43	11.0	34	1	US-08-841-636A-37
26	42.5	10.9	35	10	US-09-759-584-41
27	42.5	10.9	35	10	US-09-894-882-415
28	42.5	10.9	36	10	US-09-894-882-4
29	42	10.7	33	10	US-09-864-761-35732
30	42	10.7	34	10	US-09-864-761-40433
31	42	10.7	41	12	US-10-124-557-6
32	41.5	10.6	35	10	US-09-864-761-44585
33	41.5	10.6	44	10	US-09-879-666-7
34	41	10.5	34	10	US-09-854-864-7
35	41	10.5	38	10	US-09-864-761-41960
36	41	10.5	42	8	US-08-969-137-3
37	40.5	10.4	32	9	US-09-984-245-276
38	40.5	10.4	38	10	US-09-917-340-58
39	40.5	10.4	45	9	US-10-086-176A-1
40	40	10.2	32	9	US-09-749-637A-146
41	40	10.2	37	9	US-09-814-452-4
42	40	10.2	37	9	US-09-814-452-26
43	40	10.2	41	10	US-09-864-761-43897
44	39.5	10.1	35	9	US-10-044-359-26
45	39.5	10.1	36	10	US-09-894-882-3
46	39.5	10.1	40	9	US-10-125-459-13
47	39.5	10.1	40	9	US-10-067-761-26
48	39.5	10.1	40	10	US-09-804-156-26
49	39.5	10.1	40	10	US-09-946-633-13
50	39	10.0	36	9	US-09-916-494A-8

#### ALIGNMENTS

RESULT 1  
US-09-847-185-47  
: Sequence 47, Application US/09847185  
: Patent No. US2002076392A1  
: GENERAL INFORMATION:  
: APPLICANT: Soo Hoo, William  
: TITLE OF INVENTION: MEMBRANE-BOUND CYTOKINE COMPOSITIONS  
: COMPRISING GM-CSF AND METHODS OF MODULATING AN IMMUNE  
: RESPONSE USING SAME  
: NUMBER OF SEQUENCES: 50  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: CAMPBELL & FLORES, LLP  
: STREET: 4370 La Jolla Village Drive, Suite 700  
: CITY: San Diego  
: STATE: California  
: COUNTRY: United States  
: ZIP: 92121  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/847,185  
: FILING DATE: 01-May-2001  
: CLASSIFICATION: <Unknown>  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 09/201,931  
: FILING DATE: <Unknown>  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Campbell, Cathryn A.  
: REGISTRATION NUMBER: 31,815  
: REFERENCE/DOCKET NUMBER: P-IM 2442  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (619)535-9001  
: TELEFAX: (619)535-8949  
: INFORMATION FOR SEQ ID NO: 47:

RESULT 5  
US-09-817-647-11

```

; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Aedes aegypti
US-09-917-340-73

Query Match      12.5%; Score 49; DB 10; Length 40;
Best Local Similarity 48.0%; Pred. No. 1.2e+02;
Matches 12; Conservative 2; Mismatches 9; Indels 2; Gaps 1;

QY  42 CAPGCVCR--LGYLRRKKKVCVPRS 64
    || | : | | | | | | | :
DB  16 CAAHCIARGNRGGYCNCKKVCVCRN 40

RESULT 7
US-09-917-340-68
; Sequence 68, Application US/09917340
; Patent No. US20020090369A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Christopher J.
; APPLICANT: McAnulty, Jonathan F.
; TITLE OF INVENTION: Transplant Media
; FILE REFERENCE: TPLANT-06468
; CURRENT APPLICATION NUMBER: US/09/917,340
; CURRENT FILING DATE: 2001-07-29
; PRIOR APPLICATION NUMBER: 60/221,632
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/249,602
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/290,932
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-917-340-68

Query Match      12.3%; Score 48; DB 10; Length 40;
Best Local Similarity 29.4%; Pred. No. 1.5e+02;
Matches 10; Conservative 7; Mismatches 13; Indels 4; Gaps 1;

QY  7 GGRGKCPSEIFSRCDGRCFCPNVVPKPLCIK 40
    | | | : | | | : | :
DB  11 GNRGIC----LLNRCPGRMRQIGTCIAPRVKCCR 40

RESULT 8
US-09-865-578-4
; Sequence 4, Application US/09865578
; Patent No. US20010034433A1
; GENERAL INFORMATION:
; APPLICANT: E. MARTIN, SPENCER
; TITLE OF INVENTION: HUMAN SOMATOMEDIAN CARRIER PROTEIN SUBUNITS
; TITLE OF INVENTION: AND PROCESS FOR PRODUCING THEM; RECOMBINANT DNA MOLECULES.
; TITLE OF INVENTION: HOSTS, PROCESSES AND HUMAN SOMATOMEDIAN CARRIER PROTEIN-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 057491/0680
; CURRENT APPLICATION NUMBER: US/09/865,578
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 09/397,192
; PRIOR FILING DATE: 1999-09-16
; PRIOR APPLICATION NUMBER: 09/162,118
; PRIOR FILING DATE: 1998-09-28
; PRIOR APPLICATION NUMBER: 08/923,860
; PRIOR FILING DATE: 1997-09-03
; PRIOR APPLICATION NUMBER: 08/706,755
; PRIOR FILING DATE: 1996-09-03

```

us-09-506-978-1.slim.rapb

Thu Feb 27 09:12:56 2003

PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00663  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00662  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00661  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00670  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: US 60/234,687  
 PRIOR FILING DATE: 2000-09-21  
 PRIOR APPLICATION NUMBER: US 09/608,408  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: US 09/774,203  
 PRIOR FILING DATE: 2001-01-29  
 NUMBER OF SEQ ID NOS: 49117  
 SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
 SEQ ID NO 44588  
 LENGTH: 41  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: MAP TO AC018583.3  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.83  
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.81  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.78  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.74  
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9  
 OTHER INFORMATION: EST HUMAN HIT: A0127993.1, EVALUE 2.00e-20  
 OTHER INFORMATION: SWISSPROT HIT: Q99435, EVALUE 2.00e-21  
 US-09-864-761-44588

Query Match 11.9%; Score 46.5; DB 10; Length 41;  
 Best Local Similarity 38.1%; Pred. No. 2.1e+02;  
 Matches 8; Conservative 5; Mismatches 3; Indels 5; Gaps 2;

QY 37 LCIKICAPG--CVCRLGYLR 54  
 DB 16 MCVN--TPGSPWCICKYIR 34

RESULT 10  
 US-09-829-481-11  
 Sequence 11, Application US/09829481  
 Patent No. US20020069427A1  
 GENERAL INFORMATION:  
 APPLICANT: Presnail, James  
 APPLICANT: Wong, Zude  
 APPLICANT: Wong, James  
 TITLE OF INVENTION: Arthropod Defensins  
 FILE REFERENCE: B01441 US NA  
 CURRENT APPLICATION NUMBER: US/09/829,481  
 CURRENT FILING DATE: 2001-04-10  
 PRIOR APPLICATION NUMBER: 60/197279  
 PRIOR FILING DATE: 2000-04-14  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: Microsoft Office 97  
 SEQ ID NO 11  
 LENGTH: 37  
 TYPE: PRT  
 ORGANISM: Androctonus australis hector  
 US-09-829-481-11

Query Match 11.8%; Score 46; DB 10; Length 37;  
 Best Local Similarity 27.1%; Pred. No. 2.1e+02;  
 Matches 13; Conservative 5; Mismatches 16; Indels 14; Gaps 3;  
 QY 2 GFGGLGRGKCPSPNEIFSRCDGRCPNVVVPK-PLCIKICAPGVC 48  
 DB 1 GFG-----CPFNQ-----GACHRHCRSIRRRGGYCAGLFKQTCTC 35

RESULT 11

PRIOR APPLICATION NUMBER: 08/437,407  
 PRIOR FILING DATE: 1995-05-12  
 PRIOR APPLICATION NUMBER: 08/320,123  
 PRIOR FILING DATE: 1994-10-07  
 PRIOR APPLICATION NUMBER: 08/043,039  
 PRIOR FILING DATE: 1993-04-05  
 PRIOR APPLICATION NUMBER: 07/763,481  
 PRIOR FILING DATE: 1991-09-20  
 PRIOR APPLICATION NUMBER: 07/290,250  
 PRIOR FILING DATE: 1988-12-22  
 PRIOR APPLICATION NUMBER: 07/170,022  
 PRIOR FILING DATE: 1988-03-31  
 PRIOR APPLICATION NUMBER: 07/034,885  
 PRIOR FILING DATE: 1987-04-06  
 NUMBER OF SEQ ID NOS: 14  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 4  
 LENGTH: 42  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: MOD\_RES  
 LOCATION: (5)  
 OTHER INFORMATION: Gly or Ala  
 NAME/KEY: MOD\_RES  
 LOCATION: (14)  
 OTHER INFORMATION: Phe or Glu  
 US-09-865-578-4

Query Match 12.1%; Score 47.5; DB 10; Length 42;  
 Best Local Similarity 32.7%; Pred. No. 1.7e+02;  
 Matches 16; Conservative 2; Mismatches 14; Indels 17; Gaps 4;

QY 5 GLGGRGKCPSPNEIFSRCDGRCPNVVVPKPLCIKICA-----PGCVC 48  
 DB 6 GLGPPVRC-----XPCDARALAQ-----APPP-----AVCAELVREPCCGC 42

RESULT 9  
 US-09-864-761-44588  
 Sequence 44588, Application US/09864761  
 Patent No. US20020048763A1  
 GENERAL INFORMATION:  
 APPLICANT: Penn, Sharon G.  
 APPLICANT: Rank, David R.  
 APPLICANT: Hanzel, David K.  
 APPLICANT: Chen, Wensheng  
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
 FILE REFERENCE: A60168-X-1  
 CURRENT APPLICATION NUMBER: US/09/864,761  
 CURRENT FILING DATE: 2001-05-23  
 PRIOR APPLICATION NUMBER: US 60/180,312  
 PRIOR FILING DATE: 2000-02-04  
 PRIOR APPLICATION NUMBER: US 60/207,456  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: US 09/632,366  
 PRIOR FILING DATE: 2000-08-03  
 PRIOR APPLICATION NUMBER: GB 24263.6  
 PRIOR FILING DATE: 2000-10-04  
 PRIOR APPLICATION NUMBER: US 60/236,359  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: PCT/US01/00666  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00667  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00664  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00669  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00665  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00668

```

US-09-917-340-57
; Sequence 57, Application US/09917340
; Patent No. US20020090369A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Christopher J.
; APPLICANT: McAnulty, Jonathan F.
; APPLICANT: Reid, Ted W.
; TITLE OF INVENTION: Transplant Media
; FILE REFERENCE: TPLANT-06468
; CURRENT APPLICATION NUMBER: US/09/917,340
; CURRENT FILING DATE: 2001-07-29
; PRIOR APPLICATION NUMBER: 60/221,632
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/249,602
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/290,932
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 57
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Androctonus Australis Hector
US-09-917-340-57

Query Match 11.8%; Score 46; DB 10; Length 37;
Best Local Similarity 27.1%; Pred. No. 2.1e+02;
Matches 13; Conservative 5; Mismatches 16; Indels 14; Gaps 3;

QY 2 GFGGLGKCKPSNEIFSRCDGRCORFCPNVVPK-PLCIKICAPGCVC 48
||| ||| : : : : : : : : : : : :
Db 1 GFG-----CPFNQ-----GACHRHCRIRRRGGYCAGLFKQKTC 35

RESULT 12
US-10-136-573A-13
; Sequence 13, Application US/10136573A
; Patent No. US20020161200A1
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie Rose
; APPLICANT: Zhang, Dong Xiao
; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related Ligands and Uses Therefor
; FILE REFERENCE: P1084R1C2
; CURRENT APPLICATION NUMBER: US/10/136,573A
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 09/480,977
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: US 08/899,437
; PRIOR FILING DATE: 1997-07-24
; PRIOR APPLICATION NUMBER: US 60/052,019
; PRIOR FILING DATE: 1997-07-09
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 13
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-136-573A-13

Query Match 11.5%; Score 45; DB 9; Length 45;
Best Local Similarity 34.6%; Pred. No. 3.2e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 8; Gaps 1;

QY 35 KPLCI-----KICAPGCVCRLGY 52
| || : : ||| ||| ||
Db 10 KDFCIHGCKYVKELRAPSCICHFGY 35

RESULT 13
US-09-877-665-13
; Sequence 13, Application US/09877665
; Patent No. US20020164680A1
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie Rose
; APPLICANT: Zhang, Dong Xiao
; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related Ligands and Uses Therefor
; FILE REFERENCE: P1084R1D2C1
; CURRENT APPLICATION NUMBER: US/10/215,862
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 09/126,663
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: US 08/899,437
; PRIOR FILING DATE: 1997-07-24
; PRIOR APPLICATION NUMBER: US 60/052,019
; PRIOR FILING DATE: 1997-07-09
; NUMBER OF SEQ ID NOS: 23

```

\_\_\_\_\_

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
; FILING DATE: 16-Apr-2002
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/643,502
; FILING DATE: 18-JAN-1991
; APPLICATION NUMBER: US 07/546,114
; FILING DATE: 29-JUN-1990
; APPLICATION NUMBER: US 07/457,196
; FILING DATE: 29-DEC-1989
; APPLICATION NUMBER: US 07/390,901
; FILING DATE: 08-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Cseri, Luann
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: GI 5190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)876-1170
; TELEFAX: (617)876-5851
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
;
; US-10-124-557-8
;
; Query Match 11.3%; Score 44; DB 12; Length 40;
; Best Local Similarity 32.4%; Pred. No. 3.6e+02;
; Matches 11; Conservative 7; Mismatches 6; Indels 10; Gaps 3;
;
; QY 6 LGGRGKPSNEIFSR-----CDGRQCF---CPN 31
; Db 2 LSCGRC--PESFERGECDAQCKKDKCCPD 33
;
; RESULT 19
; US-09-814-452-5
; Sequence 5, Application US/09814452
; Publication No. US20020197689A1
; GENERAL INFORMATION:
; APPLICANT: CORZO, GERARDO
; APPLICANT: ESCOBAS, PIERRE
; TITLE OF INVENTION: INSECTICIDAL PEPTIDES AND METHODS FOR USE OF SAME
; FILE REFERENCE: 16313-0028
; CURRENT APPLICATION NUMBER: US/09/814,452
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/191,380
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Xysticus acerbus
; US-09-814-452-5
;
; Query Match 11.1%; Score 43.5; DB 9; Length 37;
; Best Local Similarity 30.4%; Pred. No. 3.7e+02;
```

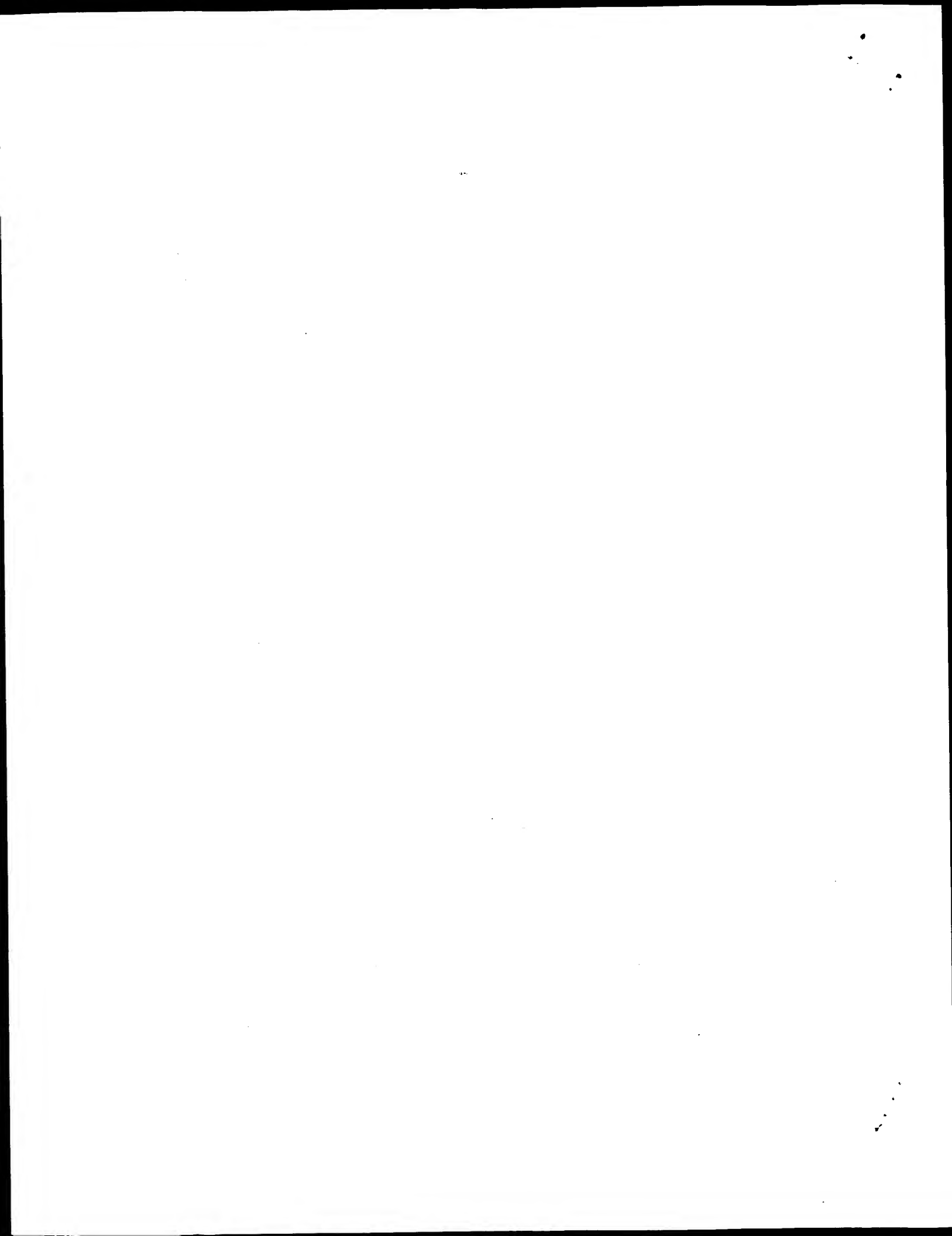
```
;
; Matches 14; Conservative 4; Mismatches 13; Indels 15; Gaps 3;
;
; QY 6 LGGRGKPSNEIFSR---CDGRQCFCPNVVVKPLCIKICAPGCVG 48
; Db 3 IGGGGGC---SVFSGPSCGGTCKCKFVLFPK-----GCHC 36
;
; RESULT 20
; US-10-136-573A-16
; Sequence 16, Application US/10136573A
; Patent No. US20020161200A1
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie Rose
; APPLICANT: Zhang, Dong Xiao
; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related Ligands and
; FILE REFERENCE: P1084R1C2
; CURRENT APPLICATION NUMBER: US/10/136,573A
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US 09/480,977
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: US 08/899,437
; PRIOR FILING DATE: 1997-07-24
; PRIOR APPLICATION NUMBER: US 60/052,019
; PRIOR FILING DATE: 1997-07-09
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 16
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-136-573A-16
;
; Query Match 11.1%; Score 43.5; DB 9; Length 45;
; Best Local Similarity 30.6%; Pred. No. 4.4e+02;
; Matches 11; Conservative 7; Mismatches 17; Indels 1; Gaps 1;
;
; QY 18 FSRCDGRQCFCPNVVVKPLCIKICAPGCVCRGLYL 53
; Db 2 FNDPCDSHTQCFPHGTCTFL-VOEDKPCACVCHGYV 36
;
; RESULT 21
; US-09-877-665-16
; Sequence 16, Application US/09877665
; Patent No. US20020164680A1
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao
; TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related
; Ligands and Uses Therefor
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/877,665
; FILING DATE: 08-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/109,206
; FILING DATE: 30-Jun-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Conley, Deirdre L.
; REGISTRATION NUMBER: 36,487
```

REFERENCE/DOCKET NUMBER: P1084R1-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: htcFalpha.egf  
LOCATION: 1-45  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-09-877-665-16  
Query Match 11.1%; Score 43.5; DB 9; Length 45;  
Best Local Similarity 30.6%; Pred. No. 4.4e+02;  
Matches 11; Conservative 7; Mismatches 17; Indels 1; Gaps 1;  
QY 18 FSRCDGRCQRCFNNVVPKPLCIKICAPGCVCRGLYL 53  
Db 2 FNDPCDSHTQCFHGTGCRFL-VQEDKPCACVCHSGYV 36  
RESULT 22  
US-10-215-862-16  
Sequence 16, Application US/10215862  
Publication No. US20030306166A1  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J.  
APPLICANT: Mark, Melanie Rose  
APPLICANT: Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related Ligands and  
TITLE OF INVENTION: Uses Therefor  
FILE REFERENCE: P1084R1D2C1  
CURRENT FILING DATE: 2002-09-24  
PRIOR FILING DATE: 1998-07-30  
PRIOR APPLICATION NUMBER: US 09/126,663  
PRIOR FILING DATE: 1998-07-30  
PRIOR APPLICATION NUMBER: US 08/899,437  
PRIOR FILING DATE: 1997-07-24  
PRIOR APPLICATION NUMBER: US 60/052,019  
PRIOR FILING DATE: 1997-07-09  
NUMBER OF SEQ ID NOS: 23  
SEQ ID NO 16  
LENGTH: 45  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-215-862-16  
Query Match 11.1%; Score 43.5; DB 9; Length 45;  
Best Local Similarity 30.6%; Pred. No. 4.4e+02;  
Matches 11; Conservative 7; Mismatches 17; Indels 1; Gaps 1;  
QY 18 FSRCDGRCQRCFNNVVPKPLCIKICAPGCVCRGLYL 53  
Db 2 FNDPCDSHTQCFHGTGCRFL-VQEDKPCACVCHSGYV 36  
RESULT 23  
US-09-817-647-16  
Sequence 16, Application US/09817647  
Patent No. US2002082229A1  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J., Mark, Melanie Rose, Zhang, Dong Xiao  
TITLE OF INVENTION: ErbB Receptor-Specific Neuregulin Related  
Ligands and Uses Therefor  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way

CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/817,647  
FILING DATE: 26-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/107,979  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Conley, Deirdre L.  
REGISTRATION NUMBER: 36,487  
REFERENCE/DOCKET NUMBER: P1084R1-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-2066  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 45 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
FEATURE:  
NAME/KEY: htcFalpha.egf  
LOCATION: 1-45  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-09-817-647-16  
Query Match 11.1%; Score 43.5; DB 10; Length 45;  
Best Local Similarity 30.6%; Pred. No. 4.4e+02;  
Matches 11; Conservative 7; Mismatches 17; Indels 1; Gaps 1;  
QY 18 FSRCDGRCQRCFNNVVPKPLCIKICAPGCVCRGLYL 53  
Db 2 FNDPCDSHTQCFHGTGCRFL-VQEDKPCACVCHSGYV 36  
RESULT 24  
US-09-030-619-197  
Sequence 197, Application US/09030619B  
Patent No. US20020035061A1  
GENERAL INFORMATION:  
APPLICANT: Krieger, Timothy J.  
APPLICANT: Taylor, Robert  
APPLICANT: Erle, Douglas  
APPLICANT: Fraser, Janet R.  
APPLICANT: West, Michael H.P.  
APPLICANT: McNicol, Patricia J.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING  
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION  
TITLE OF INVENTION: WITH ANTIBIOTICS  
FILE REFERENCE: 660081.406  
CURRENT APPLICATION NUMBER: US/09/030,619B  
CURRENT FILING DATE: 1998-02-25  
NUMBER OF SEQ ID NOS: 232  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 197  
LENGTH: 38  
TYPE: PRT  
ORGANISM: Bos taurus  
US-09-030-619-197  
Query Match 11.0%; Score 43; DB 10; Length 38;  
Best Local Similarity 33.3%; Pred. No. 4.3e+02;  
Matches 11; Conservative 6; Mismatches 8; Indels 8; Gaps 2;



Search completed: February 27, 2003, 09:11:54  
Job time : 14 secs



GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 27, 2003, 09:03:05 ; Search time 29 Seconds  
(without alignments)  
67.977 Million cell updates/sec

Title: US-09-506-978-1

Perfect score: 391

Sequence: 1 GGGFGLGRGKPCSEIFSR.....CRLGYLRNKKVCVPRSKCG 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 17865

Minimum DB seq length: 32

Maximum DB seq length: 45

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 50 summaries

Database :

- Issued\_Patents\_AA:\*
- 1: /cgn2.6/ptodata/1/1aa/5A\_COMB.pep:\*
  - 2: /cgn2.6/ptodata/1/1aa/5B\_COMB.pep:\*
  - 3: /cgn2.6/ptodata/1/1aa/6A\_COMB.pep:\*
  - 4: /cgn2.6/ptodata/1/1aa/6B\_COMB.pep:\*
  - 5: /cgn2.6/ptodata/1/1aa/PCTUS\_COMB.pep:\*
  - 6: /cgn2.6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55.5	14.2	38	2	US-08-902-516-47
2	55.5	14.2	45	3	US-08-899-437-11
3	55.5	14.2	45	4	US-09-126-121-11
4	52.5	13.4	44	6	5177197-48
5	52	13.3	42	2	US-08-761-248B-15
6	51	13.0	32	2	US-08-384-714-4
7	51	13.0	39	6	5258287-56
8	50	12.8	32	1	US-08-425-069-44
9	50	12.8	32	2	US-08-317-844B-44
10	49.5	12.7	33	1	US-08-682-485A-24
11	49.5	12.7	33	1	US-08-451-472-7
12	49.5	12.7	33	2	US-08-933-314-24
13	49	12.5	42	4	US-09-284-819-1
14	48.5	12.4	42	1	US-08-278-089A-20
15	48	12.3	40	1	US-08-033-873-10
16	48	12.3	40	2	US-08-356-832-10
17	48	12.3	40	4	US-08-988-705-10
18	47.5	12.1	43	4	US-09-230-196-19
19	47	12.0	39	2	US-08-535-494-5
20	47	12.0	39	4	US-09-097-635-5
21	47	12.0	45	1	US-08-608-786-2
22	47	12.0	45	2	US-08-824-382-2
23	46.5	11.9	38	4	US-09-384-302A-14
24	46	11.8	33	1	US-08-425-069-7
25	46	11.8	33	2	US-08-317-844B-7
26	46	11.8	37	1	US-08-212-236-10
27	45.5	11.6	43	1	US-08-179-481-81

## ALIGNMENTS

### RESULT 1

US-08-902-516-47  
; Sequence 47, Application US/08902516  
; Patent No. 5891432

; GENERAL INFORMATION:

; APPLICANT: Soo Hoo, William

; TITLE OF INVENTION: MEMBRANE-BOUND CYTOKINE COMPOSITIONS

; TITLE OF INVENTION: COMPRISING GM-CSF AND METHODS OF MODULATING AN IMMUNE  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CAMPBELL & FLORES, LLP

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: United States

; ZIP: 92121

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/902,516

; FILING DATE: 29-JUL-1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Campbell, Cathryn A.

; REGISTRATION NUMBER: 31,815

; REFERENCE/DOCKET NUMBER: P-IM 2442

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619)535-9001

; TELEFAX: (619)535-8949

; INFORMATION FOR SEQ ID NO: 47:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 38 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-902-516-47

Query Match 14.28; Score 55.5; DB 2; Length 38;

Best Local Similarity 29.8%; Pred. No. 18;

Matches 14; Conservative 0; Mismatches 14; Indels 19; Gaps 2;



ADDRESSEE: Jenkins & Gilchrist  
STREET: 1100 Louisiana, Suite 1800  
CITY: Houston  
STATE: TX  
COUNTRY: USA  
ZIP: 77002  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/761.248B  
FILING DATE: 06-DEC-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/008,348  
FILING DATE: 07-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Turley, Charles P  
REGISTRATION NUMBER: 35,723  
REFERENCE/DOCKET NUMBER: 34012.6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (713)9513310  
TELEFAX: (713)9513314  
TELEX:  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-761-248B-15

Query Match 13.3%; Score 52; DB 2; Length 42;  
Best Local Similarity 35.0%; Pred. No. 45;  
Matches 14; Conservative 2; Mismatches 20; Indels 4; Gaps 2;

QY 12 CPSNEIFSRGCRQFCFNVVPRKPLKICAPGC--VCR 49  
DB 1 CPKNP--PRSIGTCVELSGDQSCPNIOKCSNGCGHVK 38

RESULT 6  
US-08-384-714-4  
Sequence 4, Application US/08384714  
Patent No. 5840687  
GENERAL INFORMATION:  
APPLICANT: Harkins, Richard  
APPLICANT: Seto, Marian  
APPLICANT: Katz, Bradley  
TITLE OF INVENTION: MODIFIED LIGANDS FOR RECEPTOR TYROSINE  
TITLE OF INVENTION: KINASES  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wendy L. Washtien  
STREET: 15049 San Pablo Avenue  
CITY: Richmond  
STATE: California  
COUNTRY: USA  
ZIP: 94804-0099  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/384.714  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Washtien, Wendy L.

REGISTRATION NUMBER: 36,301  
REFERENCE/DOCKET NUMBER: A-0108  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 510-262-5411  
TELEFAX: 510-262-7095  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-384-714-4

Query Match 13.0%; Score 51; DB 2; Length 32;  
Best Local Similarity 50.0%; Pred. No. 44;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 38 CIKICAPGCVCRGLYL 53  
DB 6 CFHGTSPGCVCHSGYV 21

RESULT 7  
5258287-56  
Patent No. 5258287  
APPLICANT: BAXTER, ROBERT C.; WOOD, WILLIAM I.  
TITLE OF INVENTION: DNA ENCODING AND METHODS OF PRODUCTION  
OF INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN BP53  
NUMBER OF SEQUENCES: 58  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/171,623  
FILING DATE: 22-MAR-1988  
SEQ ID NO: 56:  
LENGTH: 39  
5258287-56

Query Match 13.0%; Score 51; DB 6; Length 39;  
Best Local Similarity 32.6%; Pred. No. 53;  
Matches 14; Conservative 3; Mismatches 18; Indels 8; Gaps 2;

QY 4 GGLGGRCKPFSNEIFSRGCRQFCFNVVPRKPLKICAPGC 46  
DB 5 GGLGPRVVR-----EPCDARALAQ--APPVAVCELVREPGC 39

RESULT 8  
US-08-425-069-44  
Sequence 44, Application US/08425069  
Patent No. 5728810  
GENERAL INFORMATION:  
APPLICANT: Lewis, Randolph V.  
APPLICANT: Xu, Ming  
APPLICANT: Hidman, Michael B.  
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK  
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL  
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: 301 No. 5728810th Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22046  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/425,069  
FILING DATE: 19-APR-1995  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1447-106P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX:  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: nephila clavipes  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..32  
OTHER INFORMATION: /label= silk1\_repeat  
US-08-425-069-44

Query Match 12.8%; Score 50; DB 1; Length 32;  
Best Local Similarity 80.0%; Pred. No. 56;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGGLGGRG 10  
||:|||||:|  
Db 23 GGYGGLGGQG 32

## RESULT 9

US-08-317-844B-44  
Sequence 44, Application US/08317844B  
Patent No. 5989894  
GENERAL INFORMATION:  
APPLICANT: Lewis, Randolph V.  
APPLICANT: Xu, Ming  
APPLICANT: Himan, Michael B.  
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK  
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL  
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: 301 No. 5989894th Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22046  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/317,844B  
FILING DATE: 04-OCT-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1447-105P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 241-1300  
TELEFAX: (703) 241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids

TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: nephila clavipes  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..32  
OTHER INFORMATION: /label= silk1\_repeat  
US-08-317-844B-44

Query Match 12.8%; Score 50; DB 2; Length 32;  
Best Local Similarity 80.0%; Pred. No. 56;  
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGFGGLGGRG 10  
||:|||||:|  
Db 23 GGYGGLGGQG 32

## RESULT 10

US-08-682-485A-24  
Sequence 24, Application US/08682485A  
Patent No. 5763568  
GENERAL INFORMATION:  
APPLICANT: ATKINSON, RONALD K  
APPLICANT: HOWDEN, MERLIN E.H.  
APPLICANT: TYLER, MARGARET I  
APPLICANT: VONARX, EDWARD J  
TITLE OF INVENTION: Insecticidal Toxins Derived From  
TITLE OF INVENTION: Funnel Web (Atrax or Hadronyche Spiders)  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Zeneca, Inc.  
STREET: 1200 South 47th Street  
CITY: Richmond  
STATE: California  
COUNTRY: USA  
ZIP: 94804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/682,485A  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/682,485  
FILING DATE: 17-JULY-1996  
APPLICATION NUMBER: US/08/256,933  
FILING DATE: 27-JULY-1994  
APPLICATION NUMBER: WO 93/15108  
FILING DATE: 29-JAN-1993  
APPLICATION NUMBER: AU PL0722  
FILING DATE: 31-JAN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Shaw, Melissa A.  
REGISTRATION NUMBER: 38,301  
REFERENCE/DOCKET NUMBER: PPD 5099/D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 510-231-1542  
TELEFAX: 510-231-1112  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 33 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

RESULT 12  
US-08-933-314-24  
; Sequence 24, Application US/08933314  
; Patent No. 5959182  
; GENERAL INFORMATION:  
; APPLICANT: ATKINSON, RONALD K

```

: APPLICANT: Kelly, Kathleen
: APPLICANT: The Government of the United States of America
: APPLICANT: as represented by The Secretary of the
: APPLICANT: Department of Health and Human Services
: TITLE OF INVENTION: Methods and Compositions for Inhibiting
: TITLE OF INVENTION: and Angiogenesis Comprising a Mammalian
: TITLE OF INVENTION: Subunit
: FILE REFERENCE: 015280-263100US
: CURRENT APPLICATION NUMBER: US/09/284.819

```





TELECOMMUNICATION INFORMATION

Db 3 RCSR-CTNILREPVLGCG 20

## RESULT 19

US-08-535-494-5  
; Sequence 5, Application US/08535494  
; Patent No. 5861275  
; GENERAL INFORMATION:

APPLICANT: HANSEN, J. N.  
TITLE OF INVENTION: LANTIBIOTIC MUTANTS AND CHIMERAS OF  
; ENHANCED STABILITY AND ACTIVITY, LEADER SEQUENCES  
; THEREFOR, GENES ENCODING THE SAME, AND METHODS OF  
; TITLE OF INVENTION: PRODUCING AND USING THE SAME  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT

STREET: 1755 S. JEFFERSON DAVIS HWY, SUITE 400

CITY: ARLINGTON

STATE: VA

COUNTRY: USA

ZIP: 22202

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/535,494

FILING DATE: 28-SEP-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 2747-066-0

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 39 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-535-494-5

Query Match 12.0%; Score 47; DB 2; Length 39;  
Best Local Similarity 43.8%; Pred. No. 1.4e+02;  
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 32 VVPRPLCIKICAPGV 47

: | : | | | |

Db 4 ITPQITSISLCTPGCV 19

## RESULT 20

US-09-097-635-5

; Sequence 5, Application US/09097635

; Patent No. 6153405

; GENERAL INFORMATION:

APPLICANT: HANSEN, J. N.

TITLE OF INVENTION: LANTIBIOTIC MUTANTS AND CHIMERAS OF

; ENHANCED STABILITY AND ACTIVITY, LEADER SEQUENCES

; THEREFOR, GENES ENCODING THE SAME, AND METHODS OF

; TITLE OF INVENTION: PRODUCING AND USING THE SAME

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT

STREET: 1755 S. JEFFERSON DAVIS HWY, SUITE 400

CITY: ARLINGTON

STATE: VA

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/097,635  
FILING DATE:

## CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/535,494

FILING DATE: 28-SEP-1995

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 2747-066-0

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 39 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-097-635-5

Query Match 12.0%; Score 47; DB 4; Length 39;  
Best Local Similarity 43.8%; Pred. No. 1.4e+02;  
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 32 VVPRPLCIKICAPGV 47

: | : | | | |

Db 4 ITPQITSISLCTPGCV 19

## RESULT 21

US-08-608-786-2

; Sequence 2, Application US/08608786

; Patent No. 5703049

; GENERAL INFORMATION:

APPLICANT: Rao, A. Gururaj

TITLE OF INVENTION: High Methionine Derivatives of

; Alpha-Hordothionin for Pathogen-Control

TITLE OF INVENTION: Alpha-Hordothionin for Pathogen-Control

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pioneer Hi-Bred International, Inc.

STREET: 700 Capital Square, 400 Locust Street

CITY: Des Moines

STATE: Iowa

COUNTRY: United States of America

ZIP: 50309

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/608,786

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Simon, Soma G.

REGISTRATION NUMBER: 37,444

REFERENCE/DOCKET NUMBER: 456-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 515-248-4896

TELEFAX: 515-248-4844

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 45 amino acids

TYPE: amino acid

TOPOLOGY: linear  
US-08-608-786-2

Query Match 12.0%; Score 47; DB 1; Length 45;  
Best Local Similarity 31.2%; Pred. No. 1.6e+02;  
Matches 10; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 19 SRCDGRCQFCPNVVPKPLCIKICAPGCVCRLL 50  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 2 SCCRSTMGRCMCMYMRGAMKILCAGVCRCKM 33

## RESULT 22

US-08-824-382-2  
; Sequence 2, Application US/08824382  
; Patent No. 5885802

## GENERAL INFORMATION:

APPLICANT: Rao, A. Gururaj  
TITLE OF INVENTION: High Methionine Derivatives of  
Alpha-Hordothionin  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pioneer Hi-Bred International, Inc.  
STREET: 700 Capital Square, 400 Locust Street  
CITY: Des Moines

STATE: Iowa

COUNTRY: United States of America

ZIP: 50309

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/824,382

FILING DATE:

## CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/460,440

FILING DATE:

## ATTORNEY/AGENT INFORMATION:

NAME: Simon, Soma G.

REGISTRATION NUMBER: 37,444

REFERENCE/DOCKET NUMBER: 355-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 515-248-4896

TELEFAX: 515-248-4844

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 45 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-824-382-2

Query Match 12.0%; Score 47; DB 2; Length 45;  
Best Local Similarity 31.2%; Pred. No. 1.6e+02;  
Matches 10; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 19 SRCDGRCQFCPNVVPKPLCIKICAPGCVCRLL 50  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 2 SCCRSTMGRCMCMYMRGAMKILCAGVCRCKM 33

## RESULT 23

US-09-384-302A-14

; Sequence 14, Application US/09384302A  
; Patent No. 6451543

## GENERAL INFORMATION:

APPLICANT: Kochendoerfer, Gerd G

APPLICANT: Hunter, Christie L

APPLICANT: Kent, Stephen B.H.

APPLICANT: Botti, Paolo

APPLICANT: Gryphon Sciences

; TITLE OF INVENTION: Lipid Matrix-Assisted Chemical Ligation and Synthesis  
; TITLE OF INVENTION: of Membrane Polypeptides

; FILE REFERENCE: grfn-028/02WO

; CURRENT APPLICATION NUMBER: US/09/384,302A

; PRIOR FILING DATE: 1999-08-26

; PRIOR APPLICATION NUMBER: 09/144,964

; PRIOR FILING DATE: 1998-08-31

; PRIOR APPLICATION NUMBER: 09/263,971

; PRIOR FILING DATE: 1999-03-05

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 14

; LENGTH: 38

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-384-302A-14

## Query Match

Best Local Similarity 11.9%; Score 46.5; DB 4; Length 38;

Matches 11; Conservative 3; Mismatches 13; Indels 1; Gaps 1;

QY 36 PLCIKICAPGCVCRILGRLNKKKVCVPR 63  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 12 PQCIKPKCDAGM-RFGKCMNRKCHCTPK 38

## RESULT 24

US-08-425-069-7

; Sequence 7, Application US/08425069

; Patent No. 5728810

; GENERAL INFORMATION:

APPLICANT: Lewis, Randolph V.

APPLICANT: Xu, Ming

APPLICANT: Hinman, Michael B.

TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK

TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL

TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch

STREET: 301 No. 5728810th Washington Street

CITY: Falls Church

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22046

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/425,069

FILING DATE: 19-APR-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 1447-106P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 205-8000

TELEFAX: (703) 205-8050

TELEX:

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 33 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

ORGANISM: Nephila clavipes

```
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..33
OTHER INFORMATION: /label= repeat_unit
OTHER INFORMATION: /note= "The protein of the present invention is
constituted primarily of repeats of this
sequence."
OTHER INFORMATION:
FEATURE:
NAME/KEY: Region
LOCATION: 13..18
OTHER INFORMATION: /label= alanine_stretch
OTHER INFORMATION: /note= "This run of alanine residues can also have
7 residues."
OTHER INFORMATION:
FEATURE:
NAME/KEY: Variable amino acid
LOCATION: 6
OTHER INFORMATION: /label= modified_a.a.
OTHER INFORMATION: /note= "This residue can be leucine, tyrosine or
glutamine"
FEATURE:
NAME/KEY: Variable amino acid
LOCATION: 9
OTHER INFORMATION: /label= modified_a.a.
OTHER INFORMATION: /note= "This residue can be leucine, tyrosine or
glutamine"
FEATURE:
NAME/KEY: Variable amino acid
LOCATION: 26
OTHER INFORMATION: /label= modified_a.a.
OTHER INFORMATION: /note= "This residue can be leucine, tyrosine or
glutamine"
US-08-425-069-7
Query Match 11.8%; Score 46; DB 1; Length 33;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGFGGLGGRG 10
Db 24 GXGGLGGG 33
RESULT 25
US-08-317-844B-7
Sequence 7, Application US/08317844B
Patent No. 5989894
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Xu, Ming
APPLICANT: Hinman, Michael B.
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 301 No. 5989894th Washington Street
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22046
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/317,844B
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
```

```
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 241-1300
TELEFAX: (703) 241-2848
TELEX: 248345
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Nephila clavipes
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..33
OTHER INFORMATION: /label= repeat_unit
OTHER INFORMATION: /note= "The protein of the present invention is
constituted primarily of repeats of this
sequence."
OTHER INFORMATION:
FEATURE:
NAME/KEY: Region
LOCATION: 13..18
OTHER INFORMATION: /label= alanine_stretch
OTHER INFORMATION: /note= "This run of alanine residues can also have
7 residues."
OTHER INFORMATION:
FEATURE:
NAME/KEY: Variable amino acid
LOCATION: 6
OTHER INFORMATION: /label= modified_a.a.
OTHER INFORMATION: /note= "This residue can be leucine, tyrosine or
glutamine"
FEATURE:
NAME/KEY: Variable amino acid
LOCATION: 9
OTHER INFORMATION: /label= modified_a.a.
OTHER INFORMATION: /note= "This residue can be leucine, tyrosine or
glutamine"
FEATURE:
NAME/KEY: Variable amino acid
LOCATION: 26
OTHER INFORMATION: /label= modified_a.a.
OTHER INFORMATION: /note= "This residue can be leucine, tyrosine or
glutamine"
US-08-317-844B-7
Query Match 11.8%; Score 46; DB 2; Length 33;
Best Local Similarity 80.0%; Pred. No. 1.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGFGGLGGRG 10
Db 24 GXGGLGGG 33
Search completed: February 27, 2003, 09:05:41
Job time : 30 secs
```